The Awareness and Flexibility Conjecture of Emotional Intelligence

Martin Bonus

PhD

The University of Edinburgh

2007

I hereby declare that this thesis has been composed solely by myself, contains only my own work and that no part of it has been submitted for any other degree or professional qualification.

Martin Bonus

Contents

Chapter 1: What is Emotional Intelligence (EI)?

1.1 Purpose

The purpose of this PhD was to examine a new definition of Emotional Intelligence (EI) using laboratory studies of the physiological correlates of EI processes and traditional psychometric methods (e.g., factor analysis). This multi-method approach was used to blend the strengths of experimental and psychometric designs. This thesis is specifically concerned with the 'mood regulation' aspect of EI for reasons explained below.

In this chapter background information is provided about EI, emotion, and intelligence. This material is necessarily introductory in nature as volumes could be written about any of these topics. The introduction will cover the constructs and issues of contention.

Chapters 1 through 3 will contain summaries before the main text of each to ensure the clarity of the chapter content.

1.2 Intelligence, IQ, and 'intelligences'

Because any discussion of EI must refer to intelligence theory, a very brief introduction to IQ and intelligence theory will be presented.

Mental ability testing dates back to ancient China (Jensen, 1981). More recently, Binet and Simon introduced the process of assessing individuals' performance of a series of tasks as a means of measuring their mental ability (Jensen, 1981). Tests of mental ability have been improved and adjusted over the past century to resemble what we now know as IQ tests (e.g., WAIS III-R, Ravens APM, etc). Modern intelligence tests (i.e., IQ tests) involve an assortment of tasks that have been selected to assess mental abilities primarily in the verbal and problem-solving domains. Thus most modern IQ tests have mathematical, vocabulary, reading comprehension, analogical reasoning, pattern-detection, and other sorts of items. Due to the century of debate and research that has characterised their history (see e.g., Deary, 2000; Jensen, 1980; 1981) these tests have virtually unassailable construct validity and reliability.

For the purposes of this PhD, the word 'intelligence' will refer to an individual's score on an IQ test. Of course this definition leaves out some aspects of intelligence, but as will be elucidated below, it is the most sensible and certainly the most well-supported scientific definition of intelligence. This seeming tautology will be clarified further below, but it suffices to say that the vast quantity of literature in support of IQ and the lack of any sensible alternative definition seem to justify this definition.

1.2.1 Factor Analysis and Intelligence

It is not possible to discuss intelligence without referring to the *g* factor of intelligence tests (Spearman, 1904 cited in Kline, 1996) or factor analysis. Factor analysis is a process by which the item covariances on a test are standardised and manipulated to produce a reduced number of 'factors' with corresponding eigenvectors, eigenvalues, and estimates of how much of the overall covariance is explained by each of these 'factors'. However, any 'factor solution,' or a decision of how many factors are necessary to explain the variance in a test, is almost always accompanied by 'rotation.' Factor rotation is a procedure in which the loadings of individual items on factors are forced into 'simple structure' or an approximation thereof: items are allowed to load only on one factor each. This greatly improves the interpretability of the results. As Kline (1996) and Jensen (1981) explain, factor rotation also allows us to determine whether two or more factors share variance themselves (i.e., are 'oblique' to one another and can be explained by a higher-order factor) or are uncorrelated (i.e., are 'orthogonal' to one another and should be considered as separate factors). Thus the factor analytic procedure results in a number of factors which either function separately or together to explain a set of item intercorrelations.

As Kline (1996) explains, the process of factor analysis has resulted in one of the most reliable psychological findings in existence. Simply put, intelligence tests items are positively correlated. This 'positive manifold' has been explored using factor analysis innumerable times and the reliable outcome of such analyses is that the first factor explains a great deal of the variance in all of the items. The clarity of this finding has strangely lead to a great deal of debate.

Although the positive manifold is widely agreed upon, IQ researchers argue about how many factors should be extracted from the positive manifold and the relationship between these factors. Spearman advocated that a single factor, *g*, (general intelligence) was sufficient because it explained considerably more than any other factor and because if other factors were included, rotation revealed an oblique relationship between the factors and thus that the higher-order factor, *g*, underpinned any additional factors. However, Thurstone, a contemporary of Spearman, advocated 9 'primary factors' (Kline, 1996) based on his orthogonal rotations of factors. Guilford (1967, cited in Kline, 1996) suggested that 120 factors are necessary to understand IQ scores. Gardner (1983) made a similar suggestion that there are numerous 'intelligences' which differ markedly from 'traditional' IQ, many of which can not be directly assessed. The current consensus is that the *g* factor is better split into at least two 'second-order factors' – fluid *g* and crystallised *g* (Carroll, 1993; cited in Deary, 2000). These second-order factors combined explain at least as much variance as the non-separated *g* does and also strengthen IQ theory by accounting for the variety of 'styles' of intelligence (e.g., composers, writers, mathematicians, etc). However, this debate is far from concluded. The legacy of this 'multiple intelligences' debate has been influential on EI theory.

Argument about factor structure aside, IQ tests are extremely well-researched and well-supported. Despite considerable controversy and argument about interpretation, the findings of IQ researchers are so robust as to be included in books ranging from first-year texts (e.g., Gleitman, Fridlund, & Reisberg, 2004; Smith et al., 2003) to books written for postgraduates and researchers. Experimental studies, behavioural genetics studies, correlational studies, and brain imaging studies have all suggested that IQ tests measure a 'real', meaningful construct (Kline, 1996), and as such it is probably acceptable to use it as a working definition of intelligence.

1.2.2 Mismeasurement and Mental Abilities

However, intelligence research has a chequered and contentious past. Gould (1981) critiqued the notion of mental testing on several fronts. According to Gould (1981), cases such as Sir Cyril Burt's duplicity, the errors of phrenology, and the abuses of eugenics should cause scientists and laypeople to be highly sceptical of the entire enterprise of mental testing. He also notes that the concept of intelligence testing fails because it is an example of the 'reification fallacy' in which we inappropriately assign 'real' existence out of an abstract, simplified representation of a complex series of processes. In essence, Gould's (1981) point is that the entire idea of mental testing is flawed because it is no more than an 'entity' – a score, a construct, a 'thing' – that has been created out of a complex, abstract process. The issue is that we've inaccurately assigned meaning to this 'thing' when by its very nature it can not be 'real'. In a word, just because something as vast and complex as intelligence *could* be reduced to a single score, it does not follow that it *should* be so reduced. Thus Gould (1981, pg. 239):

> The principal error, in fact, has involved...reification – in this case, the notion that such a nebulous, socially defined concept as intelligence might be identified as a 'thing' with a locus in the brain and a definite degree of heritability – and that it might be measured as a single number.

Testing is also problematic because it assumes some individuals are 'better' than others.

The reification argument should be taken seriously, especially if considered alongside the history of debate about factor analysis and the *g* factor extracted from IQ tests. If ever there was a case for this reification argument, then it should be made against factor analysis – *g* exists only in a sample's factor-analysed results and can not be directly assessed, "it is not a 'thing' with a physical reality" (Gould, 1981, pg 250). Indeed, it would not be scientifically sensible to discuss something that has no physical representation (ignoring of course the difficulty in defining 'physical representation'), so perhaps Gould is making an important claim.

As it turns out, this criticism of mental testing, or even *g*, fails for several reasons. Chief amongst these reasons is the ambiguity of the argument. It is unclear what Gould is criticising with his argument. That is, it is not certain whether it is: a) the simplification of a complex process; b) the factor analytic process itself or the seemingly-arcane g; or c) the assignment of scientific value to something that doesn't have a 'physical reality'. He does successfully attack the straw man of irresponsible methodology and the errors of past researchers, but it is not always clear what exactly he is arguing. These possible variations on the argument are taken in turn.

Possibility a) is probably accurate to some degree but it seems that Gould is being unduly harsh with IQ testing. Of course, one of the aims of science is the simplification of complex processes – the reduction of an unintelligible mass of data into succinct, interpretable reports. Indeed, all of our descriptive statistics serve exactly this purpose. Yet Gould is probably right in that simplifying complex processes can be problematic if it reaches some point of 'oversimplification.' IQ tests are certainly a simplification of a complex process but it seems inaccurate to claim that they are oversimplifications of this process. IQ tests do everything we could reasonably expect a measure of intelligence to do (Deary, 2000; Jensen, 1980;1981) and to put it simply, this definition of intelligence (simple as it might be) works extremely well. Nearly a century of research indicates as much. Perhaps the best response to a) is to point that no definition of intelligence is likely to capture *all* of the vast, complex process that is intelligence – data loss often accompanies data simplification or reduction – but it certainly captures a great deal.

Possibility b) fares little better. Gould's assertion that factor analysing IQ test results is analogous to searching for a *g* factor in "my age, the population of Mexico, the price of Swiss cheese, my pet turtle's weight, and the average distance between galaxies in the past 10 years" (Gould, 1981, pg 250) is somewhat inappropriate. The items of the former are clearly related *before* the factor analysis takes place and there is converging evidence for their interconnection (i.e., verbally astute people tend to be good at maths both in 'real life', school, and on IQ tests). We would expect intelligent people to be better at a range of cognitive activities almost by definition. In contrast, the items of the latter may be explained by a 'time' factor, but there is no *a priori* reason to suspect that they should be related to one another at all! There are indeed flaws in factor analysis, but b) only succeeds in knocking down a straw man – the spectre of poor factor analytical methodology. Kline (1996) notes that there are issues that any responsible researcher must be aware of when using factor analysis: unrotated solutions are less informative than rotated solutions; factor solutions can be subjective; and the meaningfulness of a factor analysis is limited by what is analysed. Gould does not tell us anything we did not already know.

The final interpretation, assigning scientific value or "*physical meaning*" (Gould, 1981, pg 250) to things without 'physical reality,' is perhaps more incisive as it not only applies to poor factor analyses but to the entire enterprise. However, he does not argue that *all* factor solutions lack scientific value, only *some* of them. In Gould's words (pg 250), "Sometimes this is justified…But such a claim can never arise from the mathematics alone, only from additional knowledge"(Gould, 1981, pg 250). It is unlikely that any psychometrician would disagree with this statement. It is also unclear why factor analysis of IQ test results is not justified. Surely if a test is designed to measure a single entity (i.e., IQ) then it is appropriate to use factor analysis to determine if this is the case. Gould seems to assume that intelligence can never be measured and that any attempt to simplify IQ for the purpose of measurement would be flawed, but he does not support this hypothesis particularly well. Quite simply, Gould only has only presented half an argument. His assertion that some factor analytic results are scientifically meaningless is correct but he doesn't provide any reason to believe that factor analyses of IQ tests are, especially given the amount of evidence in favour of such tests. He makes excellent points about poor methodology, the sins of previous researchers, and how our *a priori* assumptions can lead to grave errors, but in the end his work is merely based on different *a priori* belief from psychometric researchers and is not particularly compelling. A psychometric researcher works from the belief that it is important to measure constructs and compare individuals to one another because if we do not have a measure to go along with a construct, even something that stands out as clearly as intelligence would be difficult to quantify. Gould seems to think that this process 'pins down' or simplifies intelligence too much. Discussion of these viewpoints could take volumes and due to space constraints, it shall suffice to say that this thesis proceeds from a psychometric standpoint as far as intelligence is concerned.

1.2.3 Other Complaints and Responses

Gould's (1981) claim, and others like it, attack the intelligence testing process by calling into question the 'physical reality' of IQ tests. It has already been shown that Gould's claim fails, even if it is interpreted charitably, but it is possible to hamstring his claim or similar attacks at the 'physical reality' of IQ tests through other means.

Obviously the fact that IQ tests provide a paper or electronic record of a person's score confers *some* kind of 'physical reality', but there are numerous other more substantial independent physical correlates of IQ. They predict criteria we expect them to: scholastic and occupational performance, socioeconomic status, and other related criteria such as myopia (Jensen, 1981). They also fail to predict criteria we *don't* expect them to predict, such as neuroticism or agreeableness (Bastian, Burns, & Nettlebeck, 2005). There is also experimental evidence for the connection between IQ and reaction time, resting EEG potential (Kline, 1996), inspection time task performance (Deary, 2000) and specific brain activity in ERPs (Jausovec &

Jausovec, 2005). To put it simply, there is a great deal of 'physical reality' to IQ tests because the number and IQ test returns is far more meaningful than just a number. There is a wealth of converging evidence from different sources means that it is not tautological to declare that 'intelligence is what IQ tests test.' IQ tests only test verbal, spatial, mathematical, and other reasoning types yet they correlate with all of the above theoretically-linked criteria, so it is unlikely that there is a logical circle at play here. It is also worth noting that casual efficacy is not necessary for 'physical reality.' Even if IQ scores do not cause scholastic performance, the mere connection between the two still suggests that IQ scores are meaningful.

1.2.4 'Intelligences'

Gardner (1983), amongst others, has suggested that IQ alone fails to accurately describe intelligence. He argues that IQ is indeed 'an intelligence' but that it is only part of a larger number of intelligences such as kinaesthetic, musical, and inter/intrapersonal 'intelligences'. He also argues for 'interpersonal' and 'intrapersonal intelligences' which have clear similarities to the interpersonal and intrapersonal facets of EI.

Gardner (1983) lists some criteria that are necessary for a construct to be called 'an intelligence.' It must have psychometric support, be isolatable by brain damage, identifiable as a core set of operations, and have a developmental pathway, amongst other criteria. This initially seems acceptable but closer inspection reveals that Gardner's 'intelligences' often fail to meet his own criteria!

Gardner's (1983) theory lack internal consistency because his 'intelligences' do not meet some of his own criteria. For example, we would have to be pretty generous in our interpretation of 'psychometric support' if we tried to find evidence for 'kinaesthetic/bodily intelligence' or 'musical intelligence'. They can be tested (e.g., with sports ability or recital performance) but measures could hardly be considered psychometric. Worse, not only has Gardner himself not published any method of testing these 'intelligences,' but he also disavows all responsibility to do so by claiming that testing leads to labelling and stigmatisation. This view would be unacceptable to any psychometrician. Worse, his criteria lead to a sort of *reductio ad absurdum.* Personality and frontal lobe 'inhibition' both have a core set of operations, can be isolatable by brain damage, and meet his other criteria for 'an intelligence' yet they are clearly not 'intelligences.' It may be possible to make a list of things which define 'an intelligence,' but Gardner's initially promising attempt ultimately fails. Thus it is probably not scientifically rigorous to call something 'an intelligence', at least until it is clear that the criteria to make this 'intelligence' are internally consistent, psychometrically meaningful, and externally verified.

It may be that Gardner and Gould have attempted to disentangle the word intelligence from words such as 'valuable' or 'good.' This is almost certainly a valuable point to make. We live in a culture where intelligence (i.e., IQ) is rewarded but obviously in many other cultures being able to do complex maths problems (or any other task involving intelligence) would be of precious little use. Responsible psychometric researchers would not assign any value label to intelligence and this thesis certainly makes no arguments. There are multiplicity of skills which are of great use, value, and even 'goodness', and Gardner and Gould are wise to point this out, but these skills are not intelligence. A responsible intelligence researcher would probably also allow for the probability that some intelligent people do not have high IQs (e.g., people who 'don't test well') and that some high-IQ individuals are perhaps not particularly intelligent (e.g., the absent-minded professor).

1.2.5 Section Summary

It has been shown that when discussing intelligence, the best definition researchers have is one that defines intelligence as IQ test scores. This definition is not tautological because there are numerous sources of converging evidence that provide independent support for the meaningfulness of IQ tests. Some complaints have been lodged with the process of intelligence testing, but these complaints fail because they are poorly argued (e.g., they only attack a straw man) or because the body of evidence in support of IQ nullifies them.

Alternative definitions of intelligence, such as Gardner's (1984) multiple intelligence theory, are intriguing in light of 'Emotional Intelligence' but Gardner's theory fails to be internally consistent, let alone psychometrically viable. Multiple intelligence theory also fails because the criteria for 'an intelligence' do not exclude constructs which are obviously *not* intelligences (e.g., personality). His list of criteria, as well as its failure, pre-sages some theories of EI and will be discussed below.

Thus there appears to be no compelling reason to not accept the definition of intelligence used in this thesis. There is no sensible alternative to this definition and the attacks against IQ testing also seem to fail. In addition, even a brief trip to any psychology library will reveal that there is a vast body of work which supports the definition of intelligence used in this thesis. Intelligence is what IQ tests test. The debate around this subject has been heated, possibly due to the lay appeal of the construct and its social importance, but ultimately the psychometric definition seems sound.

1.3 The 'Emotion' in EI

There has also been considerable debate about what exactly emotions are and because emotions are central to any theory of EI, a very brief introduction to emotion theory is included here. This thesis is solely concerned with how to *measure* emotion, not necessarily which of the theories of emotion is most accurate. A brief introduction to the myriad theories is included because emotion is so clearly central to EI.

A physiologically oriented approach to emotion was taken for this thesis for two reasons. Firstly, although there are some EEG/ERP studies of EI (Jausovec & Jausovec, 2005) there are (at time of writing) no studies of the psychophysiological correlates of EI. Secondly, it was decided that a more 'objective' measure of emotional state might be informative in EI research which is usually assessed with self-report measures. The use of this approach neither excludes nor diminishes other approaches, it is simply the approach used here.

1.3.1 Definitional Plurality

The first notable quality of emotion theory is that it lacks anything approaching the clarity present in intelligence theory. There are literally dozens of important emotions theorists, all of whom espouse different definitions of emotion, the important facets of emotion, and even what constitutes an emotion or a theory of emotion. Even if humanistic or phenomenological theories are excluded and only psychological theories are examined, theories range from the strictly physiological or neuroscientific (e.g., Bradley, 2000; James, 1894 cited in Strongman, 2003) through the cognitive-physiological view suggested by Shachter and Singer (1962) to the 'positive psychological' books by Seligman. As Power and Dalgleish (1997) note, even within a specific type of emotion theory (e.g., cognitive theories) there are numerous sub-theories (e.g., appraisal theory, network theories) each with suggestions from different authors (e.g., the Schachter, Oatley-Laird, Lazarus, and Scherer theories). There is little common ground between these different standpoints and it is very difficult to be certain that any one is superior to the others. Thus emotion theory differs from intelligence theory in two important ways: there are many more theories and there is little consensus between theorists.

In addition to the quantity and diversity of these theories, emotions theories differ in what criteria are used to evaluate the quality of a theory. Some theorists suggest a phenomenological understanding of emotion and value ecological validity, others prefer a neuroanatomical understanding of emotion and value the seeming clarity of psychophysiological studies. In other words, emotions theories differ psychologically and epistemologically. To make matters more confusing, most theories have been supported well in journals which share their epistemological viewpoint. These differences in theory and rationale make it virtually impossible to come to a firm conclusion about which theory is 'best.'

1.3.2 From Plurality to Physiology

However, the physiological facet of an emotional response is unequivocal. Like intelligence research, the biological bases of emotion are so well-known as to be in first-year textbooks (e.g., Gleitman et al., 2004). Oatley and Jenkins (1996), amongst others, provide a succinct summary of the history of our understanding of the physiological aspect of emotion.

It is well known that the peripheral and central nervous systems are involved in emotion perception or production. The hypothalamus, thalamus, amygdala, limbic system, and the neocortex, not to mention the myriad neurotransmitters and other chemicals all have clear emotional effects (Oatley & Jenkins, 1996). Brain lesion and selective stimulation studies of animals and humans has shown that CNS activity seems to play an important causal role in emotion expression and perception whilst activity in organs innervated by the peripheral nervous system (e.g., heart, integument, pupils, muscles, etc.) seem to be reliably correlated with, if not causally responsible, for affective experience (Dawson, Schell, & Filion, 2000). The activity of these organs, also known as 'arousal,' is as reliable correlate of emotional perception or production as IQ is a correlate of academic attainment. So it would seem that it is perfectly reasonable to use physiological arousal as an indicator of affective state.

Physiological indices of emotion are not always useful for measuring complex emotions such as guilt, annoyance, jealousy, etc. Self-report indices might be better suited for these emotions. Brain imaging technology could potentially circumvent this issue, but for a number of reasons, such technology was not used in any of the studies here. An obvious workaround to this flaw with physiological indices is to simply examine simpler emotions such as general anxiety or general happiness, both of which can be easily induced using standard procedures such as affective pictures (e.g., IAPS, Lang, Bradley, & Cuthbert, 2005)

1.3.2 For the Present Purposes

As was noted previously, this thesis is only concerned with *measuring* indices of affective state, not disproving or supporting any theory of affect. The scope has so limited in order to prevent theoretical and methodological aimlessness: it is better to examine one experimental instantiation in detail than give superficial coverage to many. This thesis covers EI, not emotions theories, of which any number could be supported. It may appear blindly empirical to focus entirely on affect measurement and remain agnostic on emotion theory, but this approach is defensible.

Firstly, unlike in intelligence theory, it is defensible to separate *measurement* from *definition* when discussing emotions. This is because intelligence research examines how people differ and emotions research is more (but not solely) concerned with how people are similar. It would hardly be sensible to claim that intelligence is an ability that differs between people without some sort of measure to accompany the claim. Thus although measurement of emotion can inform emotion theory, it is not necessary to equate definition with measurement in the same way as is necessary in intelligence theory. Secondly, it is self-evident that two researchers could use identical measurement methods (e.g., self-report Likert scales) to arrive at different theories of emotion. Unlike in IQ research, measurement techniques are neither necessary nor sufficient for a definition of emotion and as such it is defensible to select a technique without accepting a specific theory. In emotions research, researchers with very different theoretical perspectives might make use of identical indices of affective state but in IQ research, the overwhelming consensus is that intelligence *is* IQ.

For the laboratory studies included in this thesis, it was decided that physiological measurement of emotion would be the most suitable. The exact definition will be discussed in detail below, but at present it suffices to say that physiological indicators of affect are of primary interest in this thesis. This measurement technique is ideal because it entirely escapes self-report measurement techniques which are endemic to many EI tests – thus it provides an index of emotional state which is entirely independent of self-report. Because self-report indices of emotion were used in most of the few experimental studies of EI there are, this is both novel and theoretically important because the results from these lab studies will inform researchers about how well EI fares when self-report methods are not used.

It should also be made abundantly clear that due to the limited scope of this thesis, it can not be argued to provide a full account of emotional experience. As has been noted, it is very difficult to arrive at a conclusive definition and so it is hoped that by taking a very specific approach it will be possible to avoid confusion and unnecessary argument. This thesis is solely focused on general affective response and will not investigate discrete or more sophisticated emotions. It is not concerned with moods and great care has been taken to avoid the phrase 'mood regulation.' As will be explained, the studies in this thesis merely involved induction of general positive and general negative affect and not any specific emotions (e.g., happiness, sadness, etc). To avoid confusion, it should also be noted that the word 'general affect' and the word 'arousal' will be used interchangeably in this thesis due to their intimate association, especially in the measures used in these studies. This reduction in scope is necessary for the goal of creating controlled laboratory studies and although the interpersonal, cognitive, and other aspects of emotional response are therefore lost, it was felt that the control was worth sacrificing 'bandwidth'. If there is any evidence in these studies for the conjecture presented in this thesis, then a logical next step would be to probe these other theories of emotion, but a limited set of goals is probably appropriate at these early stages.

1.4 In the Beginning...

Gardner's (1983) multiple intelligence theory offered two 'intelligences' which may have been the precursors to EI: intrapersonal and interpersonal 'intelligences.' These 'intelligences' refer to an individual's ability to engage with and function effectively with others and with one's self. The similarities with some branches of some EI theories are obvious.

Usually EI is attributed to an article by Salovey and Mayer (1989/90, pg.189) where it is defined as "the ability to monitor one's own and others' emotions, to discriminate among them, and to use the information to guide one's thinking and actions". Later (Mayer and Salovey, 1997 cited in Mayer, Salovey, Caruso, & Sitarenios, 2003) clarified this definition into four 'branches', or skill groups: a) emotion perception; b) use of emotions to guide thought; c) understanding emotion; and d) managing emotion. The ability to manage emotion is of particular interest in this thesis as it is the focus of all the laboratory studies here and it will be discussed in more detail below.

Two questions present themselves immediately. Firstly, because the definition of 'emotion' is such a contentious issue, using such a term in a construct seems dubious. More importantly, the 'intelligence' aspect of EI must be inspected.

As was mentioned in the discussion of Gardner's theory, we must question whether or not a set of abilities constitutes 'an intelligence.' For the sake of discussion, it is assumed that 'an intelligence' is a meaningful label that actually can be applied to a potential construct, though this is *far* from certain. Clearly, Salovey and Mayer have been influenced by the thinking of Gardner's (1983) interesting but flawed multiple intelligence theory. It seems to be assumed by Salovey and Mayer (1989/90) that what constitutes 'an intelligence' is a collection of abilities.

This definition seems to allow for obvious non-intelligences, as a simply analogy should show. 'Automobile intelligence' is (at least) the ability to depress then release the clutch while depressing the gas, the ability to judge distances and speeds, and the ability to operate to maintain a constant velocity whilst operating the transmission. Clearly, the idea of an automobile intelligence is nonsense, yet it is a collection of abilities, so by the Salovey and Mayer (1989/90) definition, it could be considered an 'intelligence.' In other words, this definition of 'an intelligence' allows for non-intelligences as well as EI and it is difficult to tell which is worthy of the designation. In an interesting but perhaps under-researched article, Locke (2005) argues that EI is flawed because it seems very broad, unrelated to intelligence, and because 'one can not reason with emotions, only about emotions' (Locke, 2005, pg 427).

Not much has been offered in response to this argument, though Mayer, Salovey, and Caruso (1999) clarify their initial definition by suggesting that not only is EI 'an intelligence' (again, assuming that this is a meaningful phrase) because it is a set of abilities, but also because these abilities are inter-correlated and share some variance with 'pre-existing intelligences' (Mayer et al., 1999, pg. 267) and because it increases with age and experience.

This refinement further reveals the influence of Gardner on EI theory but these criteria still fail. They fail to rule out the 'automobile intelligence' discussed before (because high-IQ individuals are almost certainly better drivers and it is without a doubt that driving ability increases with age and experience) and they provide another loose *reductio ad absurdum* for the entire notion of 'multiple intelligences.' By these criteria, many skill sets could be considered intelligences, including some that would normally be considered unintelligent (e.g., street fighting). In short, Mayer's attempts to distinguish a 'competency' from an 'intelligence' fail because they do not exclude ridiculous constructs. This does not imply that EI is *not* an 'intelligence' but it does suggest that any argument that EI is an 'intelligence' will need to proceed from a different angle. It could be argued that EI is related to intelligence because EI exists as some sort of orthogonal relationship to intelligence. The existence of autistic individuals with great logical abilities does seem to indicate some connection between these constructs, but this is a clinical comparison and it would be more interesting to see how the constructs would be related in the normal population.

Perhaps the crucial criterion for an 'intelligence' is that it should be almost perfectly *cognitive* in nature. This might be tacitly assumed in Mayer et al.'s (1999) work, but it certainly isn't in Gardner's (1983), which calls for kinaesthetic and musical intelligences, amongst others. If this *cognitive* criterion is combined with the other criteria Mayer et al., (1999) set out, we do arrive at something approximating a sensible definition of 'an intelligence.' EI would seem to meet these criteria as EI involves words like *thinking, monitoring, discriminating,* and other distinctly *cognitive* words. In a word, an 'intelligence' that is not cognitive is simply a set of skills and abilities. This is not to say that 'automotive intelligence' or 'kinaesthetic intelligence' are not valid collections of skills, merely that they are not *intelligence.* It is cognition that sets intelligence aside from a set of abilities.

Even with this refinement, a more conservative label than EI should probably be used. This would avoid considerable confusion and debate about the 'intelligence' facet of EI. A phrase like Emotional Competency would be more appropriate, but for the sake of convention EI is used here with the above caveats in mind.

1.5 But whose EI do we discuss?

The definition offered by Salovey & Mayer (1989/90) is not alone, although it is the only one which has claimed that EI is 'an intelligence' and suggested criteria for what 'an intelligence' is. Bar-On offered an alternative definition of EI in his earliest (EI related) peer-reviewed article as a 'Noncognitive intelligence…defined as an array of emotional, personal, and social abilities and skills that influence an individual's ability to cope effectively with environmental demands and pressures' (Bar-On, 2000, pg 1108). It is strange that Bar-On should argue that Emotional 'intelligence' is non-cognitive, but this issue is bracketed for the sake of discussion. Additionally, a multitude of definitions are available in 'popular psychology' texts (for a brief review, see Ciarrochi, Chan, & Caputi, 2000); none are worth considering in any depth as they differ little from the Bar-On definition or each other. There has been considerable debate about which theory of EI is 'the best' (see Matthews, Roberts, & Zeidner, 2002, for an introduction to this debate). The central issues of this debate are the same issues that are used to criticise EI more generally and these shall be addressed below. At present, researchers seem to have agreed to separate the existing theories into different sub-types.

Mayer et al., (1999) suggested that EI theories be split into either *mixed* theories or *ability* theories. *Mixed* theories of EI are those theories which consider EI as a collection of behavioural and personal dispositions and *ability* theories are those which consider EI a collection of specific 'emotional skills.' This distinction is not particularly convincing because it inappropriately suggests that self-report tests can be used to measure an ability. This seems awfully unlikely given that we would have little patience for self-report measures of cognitive ability or any other ability for that matter. In a word, Mayer et al., (1999) argued a theoretical split which inappropriately ignores the testing procedure used.

Such a split is inappropriate because if only a theoretical distinction is made, it is possible to arrive at psychometric absurdities. For example, there are self-report EI tests which were created to test *ability* EI (e.g., the Schutte EIS, discussed below). This doesn't seem to be a sensible approach simply because it is self-evidently dubious to measure an ability simply by asking someone his opinion of his own skills. At the risk of adding too much brevity to this discussion, the world would be full of far more intelligent people than it actually is if it were assumed that cognitive ability could be measured by asking people 'how smart are you?'.

Petrides and Furnham (2000) made a more sensible distinction based on testing methods. They argue for a theoretical and psychometric split between *trait* and *ability* EI in which the former is a collection of vaguely cognitive abilities which are tested using a collection of objective tasks and the latter is a collection of behavioural dispositions which are tested using self-report questionnaires. This distinction seems more sensible as it keeps the two theories and related psychometric practices entirely separate. To put it simply, *trait* EI is a collection of behavioural tendencies and self-beliefs measured with questionnaires whereas *ability* EI is a collection of emotion-related skills measured with purportedly objective measures. This distinction will be used in this thesis.

1.6 ***Trait* EI**

Despite the definitional issues that beset EI, there is a notable and growing body of evidence that support its meaningfulness as a construct. *Trait* EI has been especially favoured by researchers, for reasons that will be addressed below.

1.6.1 The Schutte et al. (1998) EIS

There are numerous tests of *trait* EI. Perhaps the most widely used is the un-named Schutte et al., (1998, 2001) EI scale. This measure is 33 items long and, like all *trait* EI tests, assesses EI with a Likert scale in a manner similar to personality questionnaires.

Factor analytic studies of the Schutte et al., (1998) measure have yielded equivocal results. Petrides and Furnham (2000) enumerate various reasons for caution with this measure. Chiefly, they note that the factor structure of the test is unclear: a unifactorial (overall EI) solution failed and the number of factors extracted depended too heavily on the extraction 'rule' used. Additionally, the four-factor solution they found only accounted for 40% of the variance in test responses and thus left considerable error variance. Petrides and Furnham (2000) note that there is a paucity of reverse-keyed items in this measure and thus bias from acquiescence is a potential issue. The content of the test favours a one-factor solution although the four-factor solution is tenable. Austin, Sakofske, Huang, & McKenney (2004) offer a more encouraging study. These researchers presented a 41-item variant of the Schutte et al. (1998) measure with a larger number of reverse-keyed items and showed in a large sample that both this variant and the original measure were reliable and had a reasonably clear factor structure, although they decided on a three-factor rather than a four-factor solution. Additionally, it was shown that both of these tests correlated moderately with another *trait* EI test (EQ:i, discussed below). Other researchers (Saklofske, Austin, & Minski, 2003; Gignac, Palmer, Manocha, & Stough, 2005) have found a four-factor solution, although some fit indices were less than ideal.

1.6.2 The EQ-i

This measure was developed by Bar-On (Bar-On, 1997) to measure overall EI as well as fifteen EI subscales including Self Regard, Optimism, Problem Solving, and Assertiveness. This test is owned, scored, and distributed by a corporation called MHS.

Van Der Zee and Wabeke (2004) performed one of the few independent investigations of the factor structure of the EQ-i. They used principal components analysis to extract three factors from fifteen subscale scores they analysed. Although their factor analytic procedure (PCA) and their criteria for accepting factors (eigenvalues greater than 1) are assailable, their results support what is claimed by Bar-On regarding the factor structure of this test. A similar study of the 'youth version' of the EQ-i was performed by Parker et al., (2005) in a sample of aboriginal and non-aboriginal North American youth. This test appears to have a four-factor structure rather than the three-factor structure of the normal EQ-I (Parker et al., 2005). Thus it seems that the EQ-i is a more psychometrically sound instrument than the EIS, with two important caveats. The EQ-i is not a free instrument and although MHS do provide raw item data and sub-scale data, they do not provide scoring keys for these data and thus all scoring procedures are carried out by the test distributor.

1.6.3 Other Measures

Alternatives to the longer EQ-I and the short, factorially opaque EIS have been developed, including a short-form of the EQ-i. Most notable of these are the TEIQue (Petrides & Furnham, 2003) and the SUEIT (Swinburne University Emotional Intelligence Test), both of which are the focal point of major EI research groups. There are also numerous tests which are presented and supported in single articles and have not received the quantity of attention the aforementioned tests have (e.g., Wong & Law, 2002). It is possible that these newer EI measures may address some of the limitations in the EIS.

1.6.4 Construct Validity

Many researchers have shown that *trait* EI test scores correlate with a large number of hypothetically relevant criteria. It would be expected that people who are more emotionally competent would score higher on other measures. Because emotions are central to both negative affect and social ability, it may be that high EI would provide insulation from negative affect and greater social abilities. The positive relationship between EI test scores and extroversion and agreeableness scores and negative relationship between EI scores and neuroticism scores has been confirmed many times (e.g., Bar-On et al., 2000; Dawda & Hart, 2000; O'Connor & Little, 2003; Saklofske, et al., 2003). It is likely that *trait* EI is inextricable from personality scores, and in the absence of statistical controls it is possible that the predictive value of EI is due to its relationship with personality.

Schutte et al. (1998) showed that high-EI individuals perform better at university than low-EI individuals, possibly because their EI allows them to manage the social and personal difficulties of late adolescence better. High EI seems to protect a person from some personality disorders, perhaps because high-EI seems to result in more emotional stability and more positive affect and these effects prevent clinical levels of distress (Leible & Snell, 2004). Similarly, EI also seems to predict life satisfaction (Gannon & Ranzign) perhaps because fewer bouts of negative affect result in a generally happier life. So it seems that EI tests are correlated with both personality and personality-relevant processes.

In addition to these personality-relevant correlations with EI, the convergent validity of EI tests is supported by findings that confirm that certain 'foundational' constructs are linked with EI. For example, EI is negatively associated with scores on alexithymia tests (Davies, Stankov, & Roberts, 1998; Parker, Taylor, & Bagby, 2001) which suggests that the ability to access one's own emotions co-occurs with the ability to perform high-level tasks with emotions – this is exactly what would be expected if EI is meant to be a skill set with some kind of hierarchical organisation. In a similar manner, high EI is associated with high empathy (Schutte et al., 2001) and greater affect intensity (Engelberg & Sjoberg, 2004; Dawda & Hart, 2000), which would be expected given the affective quality of these two constructs. At a 'lower' level, EI tests seem to predict performance on emotional inspection time tasks similar to those used to test IQ (Austin, 2004,2005). These IT/EI correlations, which will be discussed in detail below, have been interpreted in some cases as evidence for or against the claim that self-report EI tests measure an actual abilitiy. It has also been shown that *trait* EI scores correlate with reactivity to mood induction (Petrides & Furnham, 2003; Ciarrochi et al., 2000). These studies will be discussed in detail below but it suffices to say that *trait* EI tests have a behavioural substrate in addition to predictive validity. Also, EI is positively associated with constructs such as positive parental attachment (Kafetsios, 2004). Thus the construct validity of EI seems well supported not only because EI has important personality implications, but also because EI test responses correlate with other related constructs. Some reservations are discussed below.

1.7 *Ability* EI

1.7.1 The MEIS/MSCEIT

This is the most well-known of the *ability* EI tests. This measure began as the MEIS (Mayer EI scale) and has gone through successive iterations and alternations and is now known as the MSCEIT v2.0 (Mayer, Salovey, and Caruso EI test; Mayer, Salovey, Caruso, & Sitarenios, 2003; Mayer, Salovey, & Caruso, 2002). This test is a collection of emotion-related tasks which measure the four branches with a total of eight tasks: 1) facial emotion recognition; 2) pictures of stylised emotions and landscapes; 3) matching sensations to emotions; 4) judging which mood would facilitate which cognitive task; 5) combining two emotions to create a third; 6) deciding which emotion is made from the intensification of another emotion; 7) selecting an action to obtain a specific emotional outcome; and 8) judging which action would be best for regulating another's feelings. Performance on these tasks is measured by comparing a test-taker's responses to those of a norm and expert group.

There has been some controversy over these scoring procedures (see Matthews, Roberts, & Zeidner, 2002) and this will be discussed below.

This test and its iterations have been the subject of several detailed studies. Ciarrochi et al. (2000) indicated that the factor structure of the MEIS did not match what was predicted by the theoretical framework the test was based on. The MSCEIT, which is simply an adapted and expanded version of the MEIS, has fared better. Day and Carrol (2004) found either a two- or a four-factor solution. Similarly, Palmer, Gignac, Manocha, and Stough (2005) showed that the four-factor structure suggested in the test manual and in a recent report by the test's creators (Mayer et al., 2003) fits the data, as do other solutions.

1.7.2 Construct Validity

Ability EI test scores correlate with some of the same variables as *trait* EI scores but are more strongly correlated with measures of cognitive ability and less strongly correlated with personality test scores (Ciarrochi et al., 2000; Mayer et al., 1999; O'Connor & Little, 2003; Schulte, Ree, & Carretta, 2004; Warwick & Nettelbeck, 2004). Some studies have shown that *ability* EI tests fail to predict the same hypothetically-relevant criterion as *trait* EI tests do (e.g., Gohm, Corser, & Dalsky, 2005), but this is not surprising given that *trait* and *ability* EI tests are not strongly correlated with one another (Van Rooy & Viswesvaran, 2005; Bastian et al., 2005). There are few experimental studies of *ability* EI but it appears that performance on inspection time tasks correlates with scores on the MSCEIT in a similar manner as it does to *trait* EI tests (Farrelly & Austin, in press).

1.7.3 Section Summary

This introduction to the literature shows a flawed but maturing construct. There are at least two notable different definitions of EI and at least three major tests of these types of EI. The debate about which of these theories and which of these tests is 'the best' will be addressed below, but a brief summary of the empirical evidence presented suggest that both types of EI have some validity. If nothing else, it is clear the EI is a construct with notable predictive power and thus is worthy of consideration, despite definitional issues.

In terms of sheer numbers of studies, *trait* EI seems superior. Moreover, there appears to be more laboratory evidence of a connection between *trait* EI scores and low-level experimental tasks (Austin, 2004;2005). As noted, *trait* EI seems to predict everything it would be expected to, although this could be due to the file-drawer problem (Scargle, 2000). *Ability* EI receives slightly more support, as has been stated, but it has a smaller body of evidence in its favour.

The psychometric status of the MSCEIT is superior, with some caveats. Compared to other tests, its factor structure is clearer and more similar to what was hypothesized by the authors (Palmer et al., 2005). It has been shown that some sub-scale reliabilities of this *ability* test are low, but overall reliability is high and certainly on par with personality measures so this flaw does not reduce its relative quality as a test. McCann, Roberts, Matthews, and Zeidner (2004) suggest that 'proportion' and 'mode' scoring were preferable to consensus scoring although some advanced techniques could be used to correct for the flaws of consensus scoring (e.g., low reliability). A more serious flaw should be noted. With precious few exceptions (Palmer et al., 2005) studies have not been carried out on raw item data simply because MHS does not always provide raw item data. It is difficult to reach a scientific consensus when scientists are denied access to raw data.

Possibly the most sensible way to think of the *trait* and *ability* distinction is to examine of the component 'skills' usually attributed to EI and to discuss what each theory suggests this skill actually is. For example, emotional sensitivity (i.e., sensitivity to others' emotions) is often considered to be a facet of EI. An *ability* theorist would argue that this is a skill which should be measured by objective indices of such a skill. A *trait* theorist would argue that emotional sensitivity is a behavioural disposition. These viewpoints are probably both sensible as long as *trait* EI measures are not used to measure *ability* EI or vice versa.

1.8 'An Elusive Construct'[1]

There are profound differences between *trait* and *ability* EI theories though these theories use similar language. The debate about which is superior surrounds several issues which will be discussed.

1.8.1 Incremental Validity

As has been noted, EI tests correlate with personality and IQ tests (see Van Rooy & Viswesvaran, 2004, for a summary) and it is self-evident that the theoretical language used to describe EI is similar to that used in IQ and personality theory. Thus the incremental validity of this construct is a chief concern for many researchers.

There have been a number of studies of the incremental validity of *trait* and *ability* EI (e.g., Bastian et al., 2005; Brackett & Mayer, 2004; Ciarrochi et al., 2000;Gannon & Ranzijn, 2005;Schulte et al., 2005), most of which have shown that EI generally predicts about 5% of variance in various outcomes above and beyond personality, IQ, and demographic variables. Interpretations of these clear facts vary.

There seems to be evidence for the incremental validity of *trait* EI tests. Gannon and Ranzijn (2005) reported an R^2 change of .01 above personality and demographic variables. Bastian et al. (2005) reported an R^2 change of .05. There is similar evidence for the incremental validity of *ability* EI tests. Bastian et al. (2005) suggest that MSCEIT subscales show R^2 changes ranging from .01 to .06. Brackett and Mayer (2004) report an R^2 change of .05 and Ciarrochi et al. (2000) found an R^2 change of .04. Despite the similarity of findings between the two types of test,

[1] Davies, Stankov, & Roberts, 1998 pg 989

different writers arrive at different conclusions about whether or not either type of EI test has incremental validity over personality, IQ, and demographic variables.

It may be possible that the differences in opinion are due to the seemingly small effect size – an increase of 5% of variance explained is not large. For example, in Bastian et al (2005) the 5% is added to an R^2 of .35, for a total R^2 of .40, thus the r increases from .59 to .63. This seems small, but it is meaningful because of the shape of the r- R^2 function. That is, if we plot $f(r) = R^2$ it becomes immediately apparent that small changes in r are more meaningful (i.e., significant) as r increases. Thus a change in r from .59 to .63 is as unlikely to happen due to chance as is a change in r from .00 to .22, hence the high significance of these seemingly meaningless changes. Mathematically speaking, there is clear evidence of incremental validity.

Authors (e.g., Bastian et al., 2005) who draw the same distinction between statistical and psychological significance seem to agree with Rosenthal and Rosnow (1991), but this perspective fails to take into account the fact that small effect sizes can be observed across populations and across the lifespan. A 5% difference in predictive power may be psychologically meaningful when it is viewed across the population, across an individual's lifespan, or across the lifespan of a population (cf. the role of intelligence on cognitive aging across the population). Thus it is probably reasonable to suggest that this 5% difference is both psychologically and statistically meaningful. Also, EI appears to have incremental predictive power over demographic and other hypothetically relevant predictors (Gannon & Ranzijn), not just IQ and personality. Thus, because there is statistical and psychological evidence for incremental validity and because this incremental validity supersedes a number of predictors, it is hard to call the incremental validity of EI into question. This endorsement is irrespective of the test used because there is similar evidence for the incremental validity of *trait* and *ability* tests.

Even if it *was* the case that EI tests failed to show incremental validity over IQ and personality, which it isn't, EI could still be defended from the incremental validity complaint with an appeal to efficiency. If a researcher wished to measure EI and EI tests were excluded because of a lack of incremental validity, he/she would be forced to use a combination of IQ and personality measures instead. This seems less sensible than retaining EI which seems to represent part of the theoretical space of both of these other constructs.

1.8.2 Scoring Issues

As Matthews, Zeidner, and Roberts (2002) note, the scoring procedures used in both types of EI test are problematic. *Trait* tests are answered with a self-report Likert-type scale and *ability* tests are answered by comparing participants' responses to those of either experts or a group consensus.

These approaches all have weaknesses. Obviously assessing IQ (i.e., intelligence) through self-report would be ludicrous (for some summary and an investigation of perceived intelligence, see Furnham & Buchanan, 2005) and it is common knowledge that people over-estimate their own intelligence. Similarly, it is silly to suggest that a consensual understanding is either necessary or sufficient to make an answer 'correct' (e.g., most people don't know calculus, but it is still mathematical fact). Having experts provide correct answers may seem more sensible, but it is unclear how expert opinions on a 'sample' of simple emotional tasks could accurately represent the multiplicity of emotional situations in the real world.

The difficulty may lay in the expectations for 'objectivity' that the 'intelligence' part of the EI label confers. If EI is meant to be 'an intelligence' similar to IQ, then it should be as objective and clear-cut as IQ is. It is hard to conceive of 'an intelligence' that is tested with a scale that has 'no right or wrong answers'!
The strong form of this argument reads:

1) *Trait* EI fails as a construct because it is not objective enough

It becomes clear that this criticism is a red herring. The utility or value of the EI construct is not dependent on its objectivity. The value of the construct is dependent on its validity, the reliability of test results, and the robustness and variety of

research which supports it. Moreover there are certainly 'non-objective' self-report measures (e.g., personality) with nearly unassailable psychological value. Additionally, Austin (2004; 2005) has used an emotional inspection-time paradigm (nearly identical to the inspection time tasks used in IQ research) to show that self-report EI tests correlate with clear, low-level objective tasks. This does not mean that EI tests are objective, but it does mean that they measure something that predicts objective task performance. Thus we must discuss the meaningfulness of a construct independently of its 'objectivity' (whatever that is defined as).
Still, 1) is an understandable complaint, given the claims made by some authors and a weaker form of this attack is more threatening:

2) *Trait* EI fails as 'intelligence'. That is, if 'an intelligence' is something that is objective, it must have objective questions that are either *correct* or *incorrect*. *Trait* EI tests do not have these objective questions, therefore EI can not be 'an intelligence.'

This attack is not easily sidestepped, primarily because EI is usually advertised as some form of intelligence. It is possible that this issue would disappear if we removed 'intelligence' from EI. Obviously it is not necessary for something to be 'an intelligence' in order to be a useful construct. Also, this complaint is somewhat superficial as it condemns the construct simply because it may not meet all the expectations that a label such as 'emotional intelligence' might bring with it. Surely it is empirical investigation that should decide whether a construct is useful, not a cursory inspection of lay associations with the construct's label.

The fact that *ability* tests are scored by comparing participants' responses to those of an expert and consensual group does not mitigate claim 2). Although the MSCEIT is made of tasks rather than self-report statements, one can still question whether or not *ability* EI is 'an intelligence.'

Clearly consensual agreement is hardly necessary or sufficient for a 'correct' answer. Several more subtle points should be noted. Firstly, if 'correct' is defined as consensual agreement then the immediate implication is that an 'emotional genius' *must* appeal to everyone – a mediocre genius indeed. Few geniuses (by the psychometric definition) can even be understood, let alone appreciated, by everyone. Also if 'correct' is simply agreed-upon, then it is not possible that a person might bring external insight to the majority view – a stagnant and self-referential system is hardly 'intelligent'. This arrangement would exclude Einstein because his work did not fit in the majority view of physics! Thirdly, this type of scoring is a form of blind empiricism not unlike the MMPI that is dependent entirely on norm-group expectations which will almost certainly change as time passes – we don't really call something this malleable 'an intelligence.' Thus consensus scoring seems no more viable than self-report in light of argument 2). It may be argued that intelligence has a normative element and thus a 'common appeal' element to intelligence is merely part of its placement in the social arena, but intelligence also has a largely objective, performance-related element, so it seems sensible that intelligence should be objectively indexed.

Expert scoring (a process which results in nearly identical results according to Mayer et al., (2003)) may be the best option in light of 2). If expert scoring is used, EI tests become more like 'exams' which are marked according to what 'experts' (researchers in emotion, according to Mayer et al., 2003) know or think about emotional situations. Of course it is far from certain what exactly makes an emotional 'expert', let alone that such expertise is granted to emotions researchers. It is certainly possible that test-takers might disagree with expert views on emotional situations, but like students that disagree on interpretations of factor analysis, they might not be wrong but they are further from 'correct'. The only issue that remains is the problem of blind empiricism –scoring depends on potentially mercurial 'expert' opinions.

Because of the difficulties associated with scoring procedures for both types of EI test, it is really not reasonable to say that either is better. Although expert scoring is probably superior to self-report or consensual scoring and thus *ability* EI is probably superior in that there is more experimental research supporting *trait* EI, so it is difficult to recommend one over the other.

1.8.3 Convergent/Divergent Validity

Another contentious issue is the amount of divergent validity shown by EI tests, specifically *trait* tests. Simply put, because EI tests correlate highly with personality measures (see references above or Van Rooy & Viswesvaran, 2004, for a review) many (e.g., Matthews, Zeidner, & Roberts,2002) have argued that EI is simply 're-inventing the wheel.' Although incremental validity studies have shown this is not exactly true, it is still worth examining.

Both types of EI tests correlate with personality tests and indexes of intelligence (if not IQ itself), but the relationship between *trait* EI and IQ is weaker and less robust than that between *ability* EI and IQ. Moreover, *trait* EI tests share considerable variance with personality measures (e.g., Bastian et al., 2005). These facts have been used to argue that *trait* EI is less 'intelligence-like' than *ability* EI. There is very little in the way of a rejoinder to this critique. If something is meant to be 'an intelligence' it really ought to correlate with intelligence. The fact that *trait* EI tests do not do so has been argued to reduce their convergent validity but in fact this lack of correlation *supports* the *divergent* validity of *trait* EI. This type of EI has not been argued to be a set of skills but a collection of behavioural tendencies, so the lack of correlation is entirely acceptable or even preferable to a connection between EI test scores and cognitive ability.

However, a similar complaint can be lodged at *ability* EI. If it is meant to be 'emotional' intelligence, then it should be expected to be influenced by personality, especially neuroticism (i.e., emotionality) and extroversion. The fact that *ability* EI tests are rather weakly correlated with these measures seems to suggest that their convergent validity is also questionable.

1.8.4 Concluding Thoughts

The three topics discussed here have been focal points of the EI debate but neither theory of EI seems to be unequivocally superior. The incremental validity of the two types of EI tests seems to be roughly equivalent, both have scoring issues, and both have some issues with convergent/divergent validity.

It is fairly clear that EI is a heterogeneous construct: a*bility* EI is defined in a manner somewhat reminiscent of IQ in that it is argued to be a collection of skills and *trait* EI is defined similarly to personality in that it is a collection of behavioural tendencies and self beliefs. These similarities start in theoretical language and run through psychometric techniques to predictive power. Thus it is somewhat unreasonable to compare them directly to find the 'better' theory. This is probably what Mayer et al. (1999) had in mind when they suggested that *trait* and *ability* EI are completely different constructs.

If we view these two theories independent of the 'intelligence' label which has confused the issue, *trait* EI seems to have slightly more in its favour. The criticisms levelled at *trait* EI can be applied to *ability* EI with slight alterations and although *ability* tests have some unique strengths (e.g., the psychometric clarity of the MSCEIT), they are countered by unique weaknesses (the poor value-for-money of the MSCEIT). More importantly, there have been more studies of the predictive validity of *trait* EI by a greater number of research groups (across the world even) and there are also a greater number of freely available *trait* EI tests.

As was discussed before, the notion that EI is 'an intelligence' has some serious flaws. Simply put, it may be unreasonable to expect that EI is analogous to either IQ, personality, or both. It is probably more sensible to consider it separately from these constructs. This seems obvious, but researcher have not always ensured that hypotheses are derived from the description of the construct's constituent parts rather than the name of the construct.

1.9 EI and Affect Regulation

Both *trait* and *ability* EI include the ability to regulate moods or emotions. Because this ability is the focus of chapters 5-8 in this thesis, some space will be devoted to it.

1.9.1 What is Emotion/Mood?

As was noted above, it is not immediately clear even what 'mood' and 'emotion' refer to. Forgas and Cromer (2004), amongst others, suggest that 'moods' differ from 'emotions' in that they are long-lasting, less consciously accessible, and have a stronger effect on other cognitive processes. It is unclear why in this same study (Forgas & Cromer, 2004) they contradict themselves by using film stimuli (very 'emotion' like - consciously accessible and usually not capable of long lasting effects) to induce 'mood', but this method of discrimination may still be useful. A definition of emotion is similarly elusive and contentious, despite agreement about facial expression and emotion (e.g., Ekman & Friesen, 1986; Fridlund, 1991). Like EI, it seems that mood and emotion are aspects of behaviour that everyone understands but no one can agree upon. If it were even possible to reach a consensual definition of emotion, such a project would require volumes and might be of limited use.

The word 'affect' will be used here in place of either mood or emotion for two reasons. Firstly, the term affect is used because the methods used to induce either 'mood' or 'emotion' are nearly identical to one another, as are the self-report scales used as indices of 'mood' or 'emotion,' so methodologically it seems that a general 'affect' label is appropriate. This thesis is chiefly concerned with a very specific aspect of affect – physiological arousal. Secondly, it seems appropriate to simply refer to affect because the 'mood regulation' component of EI would probably be expected to play a role in regulation of either moods or emotions. That is, it would be expected that high-EI individuals would be better at regulating their general, unconscious mood states as well as their momentary, consciously-experienced emotions. It would seem that as far as EI is concerned, it is unimportant whether a person experiences a mood or an emotion; as long as it is affect-related, EI should play some role in its regulation.

Physiological indices of affect are used in this thesis as the primary operational definition of affective response. Specifically, skin conductance measures are used. The physiological correlates of affect are well-known, robust, and nearly consensually recognised (Bradley, 2000). It does not appear that peripheral nervous system activity is a cause of affective experience (unlike activity in some parts of the CNS) but it seems that it is a reliable correlate of affect. Amongst these PNS correlates of affect experience are heart rate increases and decreases, pupil dilation and contraction, zygomatic and other muscle activity, amongst others. According to Bradley (2000), each of these peripheral systems or organs is uniquely sensitive to either the valence (pleasure) or intensity of an affective stimulus. Skin conductance appears to be an excellent indicator of changes in stimulus intensity but is less informative about changes in stimulus valence whereas heart rate changes are strongly correlated with changes in stimulus valence but are poor indices of changes in stimulus intensity. Owing to its purely sympathetic innervation, skin conductance is a measure of flight-or-fight arousal.

The use of physiological indices of affect necessitates a simplification: the vast complexity of affect response is 'reduced' to physiological arousal. Although this issue is ameliorated in chapters 7 and 8, it is impossible to escape the fact that this thesis pays the price in ecological validity for what it gains in precision. This would be the case for any study which used physiological indices of affect but there are issues specific to skin conductance. Specifically, because skin conductance measures are good indices of the intensity of affect experience but not valence the most appropriate way to use skin conductance as a measure is to present only stimuli which are either generally positive (i.e., arousal decreasing) or generally negative (i.e., arousal increasing). Similar allowances would need to be made for other physiological indices but regardless of which indicator is used, it is clear that physiological studies of affect omit some of the subjective and experiential aspects of affect. This simplification and its importance are discussed in more detail in chapters 6 and 9.

It may also be blindly empirical for this thesis to be solely concerned with physiological arousal because such a focus does not require or entail any theory of affect. This complaint fails for several reasons. Firstly, this criticism merely knocks down a straw man – it is not claimed that affect is *only* physiological arousal (this was disproved in Schacter & Singer, 1962, if not earlier) solely that affect is operationally defined as physiological arousal for the purposes of this thesis. This claim also fails because it is simply wrong – if what is needed is a theoretical explanation, surely the text-book standard (see e.g., Gleitman, Fridlund, & Reisberg, 2004) explanation of the link between arousal and affect should suffice. Thirdly, the approach used here may be reductive, but what it lacks in application to the complexity of affect it more than makes up for in rigour and 'objectivity'. Finally, it is unclear from this complaint which of the innumerable (and often contradictory) theories of affect should be adopted rather than the present suggestion and how entertaining the present suggestion precludes other theories of emotion.

Again, it should be clarified that this thesis only examines affect. The induction techniques used and the dependant variables used to measure reaction to them all pertain to general affective state, also referred to as general arousal. The scope has been limited because it is so difficult to arrive at a certain definition of emotion. There are a multiplicity of theories of emotion, each of which could draw this thesis in a number of different theoretical directions and it would not be possible to investigate all of them in any detail. A limited scope allows greater depth and it was the goal of this thesis to examine the connection between EI test scores and measures of general affect in some detail. Care has been taken to only use the word 'emotion' where it is strictly appropriate.

1.9.2 What Is Affect Regulation?

As Gross (1999) notes, affect regulation is a difficult term to define. There are numerous issues to contend with even once a suitably specific definition is adopted.

Firstly, a number of definitions for this process could be offered. At least one is not sensible.

1) Affect regulation is the ability to control one's affect

This definition begs the question. Specifically, the word 'control' automatically implies regulation. Thus it is necessary to define this process in another manner

2) Affect regulation is the ability to change one's affect

This works slightly better but seems to lead to some new issues. It is unclear from this definition what a good 'affect regulator' changes his/her affect *to* and *under what circumstances* this person changes his/her affect. People might change their affect to be more positive or negative, but it is unclear from this definition which direction they actually choose. It is also unclear from this definition whether we are to assume that good 'affect regulators' are always in an affective state of their choosing – if they can change it at their leisure, are we to assume that if they aren't changing it, they're satisfied with it?

To address these flaws, affect regulation will be defined as

3) The ability to return one's own affective state to some 'baseline' level following any affectively positive or negative disturbance from this 'baseline' level.

This definition is not perfect. For example, the exact mechanisms involved in this process are obscure even at the level of flowchart models, let alone at the physiological level. It may also be asked if people actually live at a 'baseline' level of affect. Resolving these profound issues is beyond the scope of this PhD but given the experimental evidence in a variety of modalities for affect regulation processes (e.g., Butler et al., 2003; Gross & John, 2000) it seems that it can not be immediately dismissed.

1.9.3 EI and Affect Regulation

Affect regulation is an integral aspect of EI in most definitions (e.g., Mayer et al., 1999). It is expected that people higher in EI will be better at mood regulation. The implication for the studies in this thesis is that according to *ability* EI theory, the higher a person's EI, the better he/she will be at returning his/her affective state to some 'baseline' level. In physiological terms, it would be expected that high-EI individuals would be *less* physiologically reactive (i.e., smaller changes in arousal from baseline) to the stimuli due to their increased affect regulation abilities. As such it is critical to note that this thesis measures affect regulation *indirectly* by examining participants' changes in physiological arousal. This approach to measuring affect regulation has a precedent in the studies discussed in chapter 2. Of course, it would be ideal to measure affect regulation directly, but doing so would result in reliance on self-report measures. It would seem unproductive to use EI test questions to ask participants if they regulate their affect and then to simply repeat a similar question in a laboratory situation. However, self-report indices of affective state could be a useful supplement to physiological indicators and this combination is used in later studies in order to hone in on the relationship between EI test scores and physiological reactivity (and thus, affect regulation) from multiple 'angles.'

It also may be the case that higher EI is related to *less* effective affect regulation. Perhaps high-EI individuals are more sensitive to affective and thus react more strongly to affect induction. This thesis will examine both of these possibilities by using skin conductance activity as an indicator of affective reactivity. In these studies, reactivity and regulation are effectively inverses. Higher reactivity suggests less regulation and lower reactivity suggests more regulation, and vice versa. It should be noted that, strictly speaking, regulation is not the inverse of reactivity. Reactivity refers to how much an individual reacts to a stimulus and regulation refers to the degree to which this person alters or neutralises his affect, after the reaction has already taken place. Thus, unlike in reaction time studies, regulation is not the direct opposite of reactivity – regulation is sufficient but not necessary for reactivity. However, it seems sensible to suggest that participants' reactivity will be lower when they are specifically instructed to regulation their affect and as such, reactivity is used here as an indicator of regulation. In later chapters, self-report measures of reactivity are used alongside physiological indicators. This method indirectly assesses regulation and accomplishes one of the purposes of this thesis, which was to circumvent self-reports of EI-related behaviours.

1.10 Chapter Summary

In this chapter the definitions of emotion and intelligence to be used in this thesis were introduced and defended. It was also suggested that EI has some definitional issues related to the word 'intelligence' but that these definitional issues have been somewhat resolved by splitting EI into two sub-constructs known as *trait* and *ability* EI. These sub-constructs are theoretically different but it is differences in measurement procedures which are used to categorise the sub-constructs. *Trait* EI tests are more numerous than *ability* tests but tend to have more confusing factor structures. The nomological net of *trait* EI is better understood. The debate about which EI sub-construct has revealed a number of issues which apply similarly to both *trait* and *ability* EI tests but *trait* EI is recommend over *ability* simply because of the increased research attention paid to it, the preponderance of free tests, and the promise for the future that this attention grants. The operational definitions of affect and affect regulation that will be used for the remainder of this thesis were also introduced and defended.

Chapter 2: Experimental Investigations of EI

Most studies of EI to date have been correlational studies. The purpose of this chapter is to discuss experimental studies of EI in detail. This coverage also complements the evidence that has clarified the nomological net of EI tests. In particular, this chapter will cover studies by Ciarrochi et al. (2000), Petrides and Furnham (2003), and Austin (2004;2005), the theoretical relevance of their findings, and some specific experimental directions left unexplored.

2.1 Ciarrochi et al. (2000)

This is one of the oldest large-scale studies of EI and it is one of the first experimental studies of EI. Although this study actually involved several conjoined tasks, the discussion here will focus on their investigation of how people with different EI levels reacted to mood manipulations.

Ciarrochi et al. (2000) expected that having high EI would shield a person from the effect that an irrelevant mood might have on a judgment task. It is known that moods can influence performance on cognitive tasks (e.g., Forgas, 1998) and it was assumed in this study that participants would wish to minimize the impact that an irrelevant mood had on their task performance. The critical issue was whether or not high-EI individuals would be more capable of avoiding the influence of the irrelevant mood. That is, although everyone would want their cognitive procedures to remain unaffected by moods, only high-EI individuals might be able to ensure this was the case. The researchers also expected that the greater mood regulatory abilities associated with high EI would result in evidence of 'mood maintenance' and 'mood repair' strategies: a) high-EI individuals were expected to recall more positive memories than low-EI people after either a positive (to 'maintain' their mood) or a negative (to 'repair' their mood) mood induction; and b) high-EI individuals were expected to report more positive mood after either a positive or negative mood induction, as a result of these processes. Thus a mood x EI interaction was predicted in both the judgment task and the self-report mood state task. They also examined whether EI or IQ is a better 'insulator' from the effect of moods.

134 participants, most of whom were female, took part in a four-phase study to test these hypotheses. EI was tested with the MEIS, an *ability* EI test which was a precursor to the MSCEIT. Mood was induced first for ten minutes using a comedy programme clip (positive), a film on architecture (neutral), or a cancer death (ten minute negative). The judgment task followed this induction and after the judgment task, five minute clips were used to induce affect. After the five-minute induction, participants recalled memories from school and the positivity of their memories was used to test a different hypothesis.

The predictions were partially supported. The predicted mood x EI interaction did not occur for the judgment task, nor was there any mood main effect. There was a mood x IQ interaction, which suggested that high IQ does shield a person from the influence of irrelevant moods. There was a mood x EI interaction on the memory task: high-EI individuals recalled more positive memories in the positive and negative mood conditions than did low-EI individuals, although there were no differences in the neutral mood conditions and these effects were only observed for the *first* memory recalled after mood induction, not the second or third. Finally, it was shown that high EI individuals did report more positive affect after recalling positive memories, but not after neutral or negative memories.

The findings from this study suggest that EI effects are not always predictable and that IQ may be more useful in preventing the influence of irrelevant moods, depending on the task. The unpredictability of EI effects is something that will be observed in all of the studies reviewed in this chapter, as well as the studies in this thesis. Ciarrochi et al.'s (2000) study seems to call either EI theory or testing into question.

2.2 Petrides and Furnham (2003)

These researchers' intent was to investigate the relationship between the emotion-regulation facet of EI and performance on laboratory-based tasks. This study was in fact two experiments and because the first was concerned with face perception, a topic which is addressed again below, each will be reviewed in turn.

2.2.1 Study 1, *Trait* EI and Emotional Face Perception

Part of EI is the ability to accurately determine the emotional state of another. Petrides and Furnham (2003) predicted that because faces are so critical to emotion expression EI would include the ability to read emotions from facial displays. High-EI individuals were expected to identify facial displays of emotion more accurately (a la Ekman faces) than low-EI individuals.

The researchers split thirty four participants into high- and low-scoring EI sub-groups of equal size, based on their respective scores on the EQ-i. These participants then viewed movie-like frames of a face transforming from a neutral face to one of the 'basic' emotions (Ekman & Friesen, 1996): three 'films' each for anger, sadness, surprise, happiness, disgust, or fear. These 'film' clips were actually still images arranged in order that presented each still image for two seconds before moving to the next automatically, as in an automated slide presentation. Each emotional 'film' was comprised of 21 still images. The mean latency for identification of each emotional expression as well as the number of images required for identification was recorded.

The hypothesised difference in ability to identify emotional expressions was supported, with high-EI individuals recognising emotions faster and (necessarily) with fewer stills. Thus this study provides support for EI theory by showing that in lab tasks, EI differences correspond to task performance differences. There were perhaps issues which could be improved in this study: the EQ-i does not contain any facial emotion recognition items and the task itself was fairly unchallenging, even for low-EI individuals. These types of flaws are hardly insurmountable and a replication that used a different EI test or perhaps a slightly more difficult task (perhaps using faces for 'guilt' or other complex emotions) could be easily done. As Petrides and Furnham (2003) note, the fact that the task performance differences exist even with a limited test such as the EQ-i only serves to support the suggestion that EI tests to seem to test some kind of emotional competence. As an aside, it would be interesting to determine if performance on this task correlated with performance on the facial perception component of the MSCEIT, as the two tasks seem very similar. It may even be the case that the task used here could be adapted as an alternative to the MSCEIT.

2.2.2 Study 2, Affect Regulation

Their second experiment is of particular note for this thesis. Petrides and Furnham (2003) investigated affect regulation using a different logic to arrive at an opposite hypothesis from that of Ciarrochi et al. (2000). Based on their first study which showed that higher EI is associated with more accurate emotion detection, they predicted that high-EI individuals would be more 'susceptible to affect' (Petrides & Furnham, 2003, pg 46). That is, it was predicted that EI test scores would be associated with greater reactivity to affect induction (cf. chapters 5-8 of this thesis, Ciarrochi, 2000).This is an interesting prediction to make because EI is not only characterised by greater emotional sensitivity, but also greater emotional control. It is hard to know which would be visible in behaviour! Here, Petrides and Furnham (2003) highlight one of the key difficulties in EI research: in a word, hypotheses can go in many directions with equally compelling logic on each. It should be noted that this is an issue which does not apply to *trait* EI because this form of EI allows for people to hold mutually exclusive self-beliefs or engage in contradictory behaviours. The issue of 'conflicting abilities' is more of a dilemma for *ability* theory.

In this study, high- and low-scoring groups of fifteen were made from previous participants who were re-contacted. The researchers also residualised the participants' TEIQue scores for personality differences so the final ranking of participants from high to low had any personality variance controlled for. In a mood induction paradigm, participants completed an emotional status questionnaire upon arrival at a lab (baseline), after viewing a distressing film (negative), and after viewing a comedic film (positive). Responses to this questionnaire can be summed for a tension, depression, anger, vigour, confusion, and overall general negative affect score.

The prediction of greater impact of affect induction was supported in all of the emotional state scales (e.g., anger, tension, etc) with the exception of 'confusion.' In addition to this overall tendency toward greater reactivity, it appears that high-EI individuals are more affected by *changes* in mood stimuli than low-EI individuals are: the negative film clip seemed to distress them more and the positive film clip seemed to 'restore' more positive affect than low-EI controls.

It is difficult to fathom why Petrides and Furnham (2003) found a relationship where Ciarrochi et al. (2000) found none. It may be the case that these differences are due to differences in experimental procedure: Ciarrochi et al. (2000) used a more complex method of measuring mood reactivity – through judgment tasks and memories – than did the other researchers. It also may be the case that *trait* EI, examined in Petrides and Furnham's (2003) article, predicts behaviour in a different manner than does *ability* EI. It could also be the case that the discrepancy in findings is due to the residualisation process that Petrides and Furnham (2003) used to control for personality variance – their findings may more accurately reflect how EI connects to behaviour than other studies.

Because these two studies yielded different findings implications for EI theory, a replication would be enlightening. It would be a simple process to examine the residualised scores and compare them to EI results that had not been residualised, in order to determine if the residualisation process was responsible for these findings. Other methodological differences could be ruled out by a series of replications. The more important theoretical questions are less easily resolved.

It is not immediately apparent why 'sensitivity' seems to win out over 'control' in this experiment or indeed why one facet of EI would override another in any experiment. It may be the case that if EI is a hierarchical construct (e.g., Mayer, 1999, but see chapter 1 for issues with factor structure of EI tests), the 'lower-level' abilities such as emotion recognition are dominant in the absence of any requirement for 'higher-level' activities. That is, perhaps the 'sensitivity' response wins out over 'control' if there are no requirements for such a 'high-level' ability. In Ciarrochi et al.'s (2000) study, participants had a cognitive task to perform after mood induction (the judgment task) whereas in Petrides and Furnham's (2003), there was no such task, so there was no real need for any kind of 'control'. An experiment which manipulated whether or not high-EI individuals were told they would need to 'control' would determine if this is the case and such an experiment is carried out in Chapters 6 and 7. Alternatively, the dominance of 'sensitivity' could be due to it being more 'instinctual' than 'control' – perhaps it is an automatic response that must be actively suppressed in favour of 'control'. It may also be the case that *trait* EI scores are positively associated with emotional sensitivity whilst *ability* EI scores are positively associated with emotional control. However, these possibilities fail to account for Ciarrochi et al's (2000) inability to reject the null hypothesis in some cases and it is hard to substantiate some of these suggestions without any experimentation. Part of the purpose of this thesis was to address the paucity of experimental studies of EI and although it has been geared to answer a specific question about a specific conjecture about EI, chapters 5-8 also serve to replicate findings about affect regulation and EI.

It may also be argued that the potential 'conflict' between sensitivity and control is more of an issue for *ability* EI than it is for *trait* EI. Because *trait* EI is a collection of behavioural dispositions and self-beliefs, it is perfectly reasonable to predict that people's beliefs about and tendencies towards *both* control and sensitivity would correlate with objective indices of their levels of control and sensitivity. On the other hand, an *ability* theorist would have to explain why it is not clear *a priori* which of these two skills will 'take control' in a given experiment.

2.3 Austin (2004,2005) and Farrelly and Austin (in press)

Austin (2004,2005) has carried out a number of experiments which have examined the correlation between EI score and performance on an emotional face inspection time task. Although her methodology differs from that of Petrides and Furnham (2003), she has found similar support for EI by showing that higher *trait* and *ability* EI is associated with better performance on emotion-relevant tasks.

2.3.1 Austin, 2004

The prediction for this study is simple and is analogous to the predictions made about the correlation between IQ and reaction/inspection time. Simply, Austin (2004) predicted that higher EI scores would be associated with better performance on an emotional inspection time task.

Undergraduates and members of a more mature volunteer panel completed the NART, a personality measure, the adapted version of the EIS (Austin et al., 2004), and completed three experimental tasks: a face inspection time (IT) task (either sad or happy), a control symbol (+ x) inspection time task, and an untimed Ekman emotion recognition task. The emotional inspection time task differed from the Ekman task in several important ways. First, each trial had a face-neutral mask-prompt form, whereas the Ekman task had a single face with a multiple choice prompt. Secondly, the IT task was timed with the emotional face presented for a fixed duration no longer than 350 ms.
The hypothesised correlation between EI and IT performance was partially supported. The 'appraisal' aspect of EI, arguably the most immediately relevant to emotion perception, was significantly positively correlated with performance on the two emotional IT tasks and the correlation between this EI sub-scale and the Ekman task approached significance. Because the IT tasks were intercorrelated, the variance attributable to the Symbol IT (control) task was partialled out and the correlation between EI and Sad IT remained.

The results of this study suggest that *trait* EI does correlate with performance on objective emotional tasks in the lab, albeit not very strongly. This correlation remains after controlling for participants' general IT ability. It is strange that Austin (2004) observed a significant correlation *only* for those EI items that measure emotional appraisal when Petrides and Furnham (2003) found similar results *without any* appraisal items. Still, these findings suggest that *trait* EI measures something 'real.' They also suggest that *trait* EI is a useful way of measuring objective skills because it is clear that people who *say* they can perceive emotions well clearly *do* perceive emotions well.

2.3.2 Austin, 2005

This study was a replication of the previous study with several variations. In addition to the Ekman face task and the sad/happy IT task from Austin (2004), this study included a word inspection task. This word inspection task was actually two tasks: to decide whether a word is emotional or non-emotional and whether a letter string is a word or a pronounceable non-word. Also, the EQ-i short form was used to measure EI as well as a shorted version of the adapted (Austin et al., 2004) EIS and the Raven's Standard Progressive Matrices was used to test fluid ability. Participants were all drawn from a volunteer panel of older individuals.

The results from this study were similar to those of the previous study although the EI correlations were less robust. The 'interpersonal' facet of the revised EIS correlated with performance on the combined emotional IT task. It does not appear that this correlation was calculated after controlling for general IT ability as was done previously.

This article appears to be further evidence for the predictive validity of *trait* EI tests in the lab. However, some important caveats to this statement must be noted. First, the correlation between EI test scores and IT performance was weaker than in the previous study (Austin, 2004). This could be due to the age of the sample because the prior study had been carried out largely on young people. It could also be due to the EI test that was used, but this would be strange in light of Petrides and Furnham's (2003) finding that the EQ-i predicts emotional face perception ability. It must be asked why only a specific aspect of EI correlates with emotional IT performance if emotion perception is so fundamental to EI. To put it baldly, there is far too much variation between these experimental studies to allow any kind of consensus, except possibly on the finding that 'interpersonal/appraisal' EI correlates with emotional IT performance. These IT-EI studies have not been widely replicated. It may seem that the evidence does not outweigh the probability of a simple Type 1 error but the specific nature of the findings above seems to hamstring this argument. Also, it may not be justified to expect that *all* EI sub-scales correlate with IT task performance – not all IQ tests correlate with IT tasks either (e.g., the NART, Austin, 2004). In all, these studies provide some introductory evidence in favour of the construct validity of *trait* EI and although the support must be taken with the proverbial pinch of salt, it does appear that *trait* EI tests do index an objective ability.

2.3.3 Farrelly & Austin, in press

This article was similar to the prior studies (Austin, 2004;2005) but also addressed the dearth of experimental studies of *ability* EI in the lab. A correlation between MSCEIT score and emotional inspection time task performance is especially relevant because the MSCEIT itself contains tasks that are similar to the emotional IT task and because MSCEIT scores tend to correlate with IQ test scores, which are correlated with IT performance. Farrelly and Austin (in press) also predicted that the two types of EI would correlate with one another and that IQ would correlate with scores on the MSCEIT.

The first experiment in this study examined the correlation between scores on *ability* (MSCEIT) and *trait* (Austin et al., 2004 variant) EI, intelligence (measured with the Gf/Gc Quickie Test Battery) performance on the IPT-15, a task that measures the ability to detect lying and social roles based on verbal and non-verbal cues, and performance on the emotional face and neutral symbols tasks.

Some of the hypotheses were supported in this study. Scores on the *ability* and *trait* EI measures correlated positively with one another as was expected, but scores on the IQ tests did not correlate with MSCEIT scores. The lack of correlation between IQ and *ability* EI seems at odds with the predictions and findings of Mayer et al. (1999). MSCEIT scores did not correlate with any of the emotional IT tasks, but they did correlate positively with performance on the IPT-15 and, counter-intuitively, they correlated negatively with neutral symbol IT performance. *Trait* EI test scores correlated with IPT-15 but not the emotional IT task.

These results contradicted past results (e.g., Austin, 2004;2005) in studies of both types of EI test. The second experiment was performed to determine if the findings of the first experiment were anomalous. The second experiment used more participants, a different measure of fluid *g* (Raven's APM), the Ekman faces task in addition to the IPT-15, and the EQ-i short-form instead of the Austin et al. (2004) variant of the EIS. Also, the happy face portion of the IT task was omitted but the sad face and symbol parts were retained.

The results from this study supported the hypotheses to a greater extent than the those of the first study. Scores on the MSCEIT correlated positively with crystallised *g* but not fluid *g*, the Ekman faces task but not the IPT-15, and the sad IT task but not the symbol IT task. Scores on the MSCEIT and the EQ-i were correlated positively but EQ-i scores were not correlated with any of the lab tasks.

The findings from these two studies seem to suggest that the MSCEIT measures an actual ability that is recognisable in similar laboratory tasks. The results of the second experiment seem to suggest that the MSCEIT predicts objective task performance better than *trait* EI tests, although the results from the first experiment differ. However, these reasonably clear findings come packaged with some odd findings. The correlation between MSCEIT and the IPT-15 was not replicated in the second experiment and it is suggested that low reliability could have caused this. However, it is unclear why a task with such low reliability was used in the analyses at all. Additionally, there were unexpected patterns of correlations between the measures of fluid *g* and the tests and tasks in both studies. Despite claims to the contrary, ability EI tests scores tend to not correlate with measures of fluid ability (e.g., Ciarrochi et al., 2000), so perhaps these correlations are understandable. The finding that performance on the emotional IT task did not correlate with fluid ability initially seems problematic for the paradigm used in these studies, but because the correlation between IT tasks and fluid ability varies from study to study (Deary, 2000), Scores on the *trait* EI tests failed to predict any of the lab tasks, but this is hardly surprising given that overall EI did not predict task performance in either of the prior studies (Austin, 2004;2005) either. The finding that performance on only part of some *ability* EI tests correlates with IT task performance seems especially anaemic in light of the robust correlations between the MSCEIT and these same tasks.

It also may be the case that the connection between *ability* EI test scores and IQ is simply due to the psychometric overlap between the two constructs. As *ability* theorists are quick to point out, this type of EI is theoretically and psychometrically related to IQ. It may be the case that all that the correlation between MSCEIT and IQ scores demonstrates is that some MSCEIT processes are similar to IQ test processes. Perhaps this correlation merely demonstrates that content overlap leads to psychometric overlap. This possibility might not trouble an *ability* theorist, however, as he might simply reply by saying that the purpose of the MSCEIT was in fact to test EI like IQ is tested. Also, this criticism could be levelled at the *trait* theorist as well: perhaps the association between personality and *trait* EI is merely due to a similarity in how the questions are asked.

2.3.4 Concluding Thoughts about Austin (2004;2005) and Farrelly & Austin (in press)

It is difficult to be certain what these experiments show us as there are some issues that cloud the otherwise clear results. The IT task, at least in the most recent article, does not appear to correlate with fluid ability in a manner we'd expect an IT task to. This may be considered acceptable given that this IT task was designed to be about very basic emotion perception, but given that Farrelly and Austin (in press) themselves note that inspection time tasks are usually correlated with fluid ability, the lack of any similar correlation is somewhat disconcerting. Perhaps a replication with a different IT task may yield similar results, although Burns and Caryl (manuscript in preparation) have found null correlations between *trait* EI and IT task performance when using a different task. There is some further confusion as the findings do not consistently replicate across studies and it is not immediately clear whether this is due to the low reliability of some tasks as well.

These studies still shed a great deal of light on the construct validity of EI. They effectively doubled the experimental EI literature, presented a novel experimental paradigm, investigated both *trait* and *ability* EI, and also replicated/challenged findings about IQ and EI. In addition to these qualities, they also indicate that both *trait* and *ability* EI tests seem to measure something that can be independently measured with experimental tasks, even if one might disagree with the task used here. These studies also suggest that when the conditions are right, the MSCEIT is a better predictor of lab-based tasks.

2.4 Interpretation and the Relevance of Experimental Studies to Theory

These studies are important because some of them allow some preliminary causal inferences to be draw and because they shed light on how objective abilities (e.g., facial display perception, affect regulation) correspond with different levels of *trait* and *ability* EI – this would not have been possible without experimentation.

Ciarrochi et al.'s (2000) and Petrides and Furnham's (2003) studies showed that high- and low-EI sub-groups react differently to mood induction from one another. This is similar to the lab studies of extroversion and physiological arousal. Although these two studies yielded different results, it is certainly clear that there are meaningful differences in emotional sensitivity and emotional regulation between high- and low-EI individuals. These are precisely the results that would be expected according to EI theory and the fact that they were found in different labs with different tests stands as convincing if preliminary evidence that EI tests measure the abilities they is advertised to measure. The results of Austin's (2004;2005) and Farrelly and Austin's (2005) studies are also very informative. Their facial perception task results suggest that there are basic tasks at which high-EI individuals outperform low-EI individuals. Again, these findings support the claims that EI theorists make. These studies also used laboratory tasks with pre-defined correct and incorrect answers which allowed for greater precision in interpreting results compared to the self-report methodology used in most questionnaire studies. Although participants' EI levels were not manipulated such that it is possible to be certain that EI differences cause skill differences, the different 'experimental' groups still perform differently and as such it is possible to draw preliminary causal inferences.

These experimental studies are especially informative because they bypass the correlation-causation problem that is endemic to correlational studies. Although (see chapter 1) it certainly appears that EI tests measure something that correlates with many hypothetically predicted criteria, the wide range of the correlations and the overlap with other important variables (e.g., personality) make it hard to rule out the possibility that EI is caused by the criteria or a third variable. It is easier to argue that higher EI causes better task performance in such experimental studies as the samples are presumably random and thus not meaningfully different from one another, save in EI levels. Also, it is hard to think of a way that task performance which occurs *after* an EI test could affect EI test scores. Of course, it is possible that a third variable is causing the EI-task link given that EI tests correlate with other variables, but based on Austin's (2004;2005) and Farrelly and Austin's (2005) work, it probably isn't crystallised *g*, and it certainly isn't fluid *g*. Future studies could be designed to include other EI correlates such as personality in order to ensure that only EI is used to separate participants into high/low sub-groups. Yet although these studies leave many questions unanswered, they seem to suggest that when EI differs, behaviour on a number of objective, lab-based, emotion-related tasks changes as well. This is precisely what would be predicted according to EI theory and these lab findings greatly augment the questionnaire results reviewed in chapter 1.

2.5 Future Directions

In addition to some refinements to the IT studies, it would be interesting to replicate and expand upon the affect regulation results obtained by Ciarrochi et al. (2000) and Petrides and Furnham (2003). As was noted above, the findings of these two studies were opposed, which is interesting in light of an EI theory that suggests a unitary construct – it is unclear how to reconcile the simultaneous, opposing action of two constituent parts of a single unitary construct. Additionally, there are some methodological changes that could be made to improve our understanding of the construct.

Specifically, both of the affect regulation studies discussed previously relied on self-report methodology to assess emotional state. There has not, at present, been a study of EI and affect regulation that has used an objective index of emotional state. This is more important than is initially apparent. Because both of the prior studies of EI and affect regulation have used self-report methodology to assess mood, it is possible that some sort of self-report response bias has contaminated the prior studies. It is possible, even plausible, that people who score highly on self-report EI questionnaires, specifically on items that measure mood regulation (e.g., 'I am good at controlling my mood'), would self-report their own emotions in a way that confirms their own self-perceptions. This may be an unavoidable consequence of self-report methodology and it was hoped that using physiological indices of affect in this thesis would circumvent this issue. In other words, it may be the case that any observed link between high EI and better affect regulation is simply due to people fibbing or reporting whichever emotional state that would prevent cognitive dissonance, probably not even intentionally. The use of an objective index of affective state would allow researchers to circumvent this issue because such an index would be independent of a participants' self-evaluation of EI level or affect.

Although this is covered in greater detail in chapter 5, a number of 'objective' measures of affective state exist. Psychophysiological measures such as skin conductance response, pupil dilation, heart rate, and zygomatic eye muscle activity (amongst others) each have unique advantages and disadvantages (Bradley, 2000) and would all would be suitable here. They are all presumably independent of one's self-reported affective state in that most people are probably unlikely to alter their skin conductance simply for the purpose of fooling an EI researcher. In other words it seems unlikely that these indices would suffer from self-report bias because they are not self-report measures. Also, as was noted in chapter 1, these indices are independent of the theory of emotion a researcher endorses, so it may be easier to interpret studies that examine the psychophysiological correlates of EI without the theoretical 'baggage' attached to theories of emotion. All of the experimental studies in this thesis use psychophysiological methodology.

A psychophysiological study of EI might also be useful in settling the aforementioned confusion about why some EI-related skills seem to override others. It was suggested that this discrepancy was due to one study involving a 'high level' task that required clear thinking and the other only involving self-report ratings. More attention is given to this possibility in chapter 3. It would be straightforward to design a study which manipulated how explicit the affect-regulation requirements were to participants and such a study would help determine what conditions result in an increased role of which EI-related skill. It is exactly this kind of experiment which is the subject of chapters 6 through 8.

2.6 Chapter Summary

In this chapter, a number of experimental studies of *trait* and *ability* EI were examined in detail. Although the results were in some cases contradictory and in other cases made unclear by additional questions, there can be no mistaking the finding that EI is linked to objective indicators of affective state. These studies also suggest that *ability* EI test scores, under the right circumstances, correlate more reliably with task performance than *trait* EI test scores. The value of these studies was highlighted and suggestions for their improvement were made. Special attention was given to the possibility of similar lab studies which used physiological methods instead of self-report indices of mood regulation. At this point, current EI theory and those EI studies of note have been covered in some detail, and in the next chapter, the *awareness* and *flexibility* conjecture will be introduced.

Chapter 3: The *Awareness* and *Flexibility* Conjecture

3.1 Introduction

The previous chapters have introduced some of the antecedents of EI and some of the essential issues surrounding the construct. Several important experimental studies of EI have also been discussed in detail. Through this discussion it has become clear that EI is a construct beset with some definitional difficulties that it seems to be increasingly well supported in the literature in spite of these difficulties. In this chapter, it is suggested that current EI theory omits an important piece of theory. This conjecture will be presented, critiqued, and examined for hypothetical implications.

3.2 The Hole in Current Theory

Present theory suggests that EI-related skills (e.g., facial display detection, affect regulation, etc) are on a continuum that relates to the EI test score continuum in a simple, linear manner. That is, high-EI individuals are thought to be better that affect regulation than low-EI individuals, just as high-IQ individuals tend to be better at maths than low-IQ individuals.

This is perfectly reasonable but it fails to account for the ambiguity that characterises emotional tasks. The tasks and abilities that comprise IQ have obviously correct and incorrect answers and because of this it is possible to say that higher IQ leads to better (i.e., more correct) performance. Emotional tasks and abilities lack these kinds of clear, right/wrong answers and as such it is not possible to assume that higher EI will necessarily relate to 'better' performance. It is likely that 'better' performance on an emotional task would be a different, perhaps even contradictory, response in different situations. A confrontation with a loved one provides an anecdotal example: in some cases it may be 'better' to pursue the matter and settle it immediately, in some cases it may be 'better' to let the issue wait for a future date –

it depends on the actors and the situation. In a word, being right/wrong depends entirely on the situation, and current EI theory does not take this into account.

It is possible that EI tests may not relate in a simple manner to performance. In the example above, it is possible that the same high-EI individual might respond in an entirely different manner given two subtly different situations. Or two equally high-EI individuals might respond differently to the same item simply because they 'read' the item in subtly different manners. Both solutions are probably correct, but present EI theory does not account for this possibility.

3.3 The *Awareness* and *Flexibility* Conjecture

What is proposed for this thesis is an addition to current EI theory. It should be made clear from the beginning that this conjecture does not replace current theory, it merely augments it. In fact, one branch of this conjecture clearly works hand-in-hand with the skills associated with current EI theories. Simply put, the conjecture is that EI is greater *awareness* and *flexibility*. *Awareness* refers to sensitivity to the intricacies of emotion-related 'problems' and the variety of manners in which one can approach these dilemmas. *Flexibility* refers to the ability to actualise these varieties of solutions. *Awareness* would be a necessary, but not sufficient condition for *flexibility* and the latter would be the more sophisticated of the two components. It is worth discussing these in greater detail before moving on to discuss experimental implications.

It may seem that *awareness* is already included in present EI theory as the 'emotional perception' facet of EI, but this is not the case, although the two are complementary. *Awareness* refers to a very specific type of emotional perception: recognition that there are often multiple solutions to a problem and only specific situational constraints can illuminate which is the ideal solution. It may seem obvious that many emotional 'problems' have multiple solutions, but anecdotally, most people could probably list people they knew who reliably take the same approach to interpersonal challenges, regardless of who is involved or what the specific situation is. And constructs such as the autism spectrum (non-clinical; Baron-Cohen, Wheelright, Skinner, Martin, & Clubley, 2001) suggest that some people don't think about even the most simple emotional ramifications of interpersonal relations, let alone a variety of approaches to solve an interpersonal dilemma. Of course, *awareness* and simple emotion perception are intricately linked; *awareness* is actually just a specific type of emotion perception that is not currently part of EI theory. Thus this conjecture simply adds an incremental theoretical step to EI theory.

Flexibility has a small amount of conceptual overlap with current EI theory but is the more important, measurable aspect of this conjecture. There is anecdotal support for this addition to EI theory in that everybody probably knows people who always take the same (usually extreme) interpersonal approach to all problems, regardless of who is involved or the subtleties of the situation. There have also been several decades of personality research has shown that some people are optimistic, pessimistic, aggressive, passive, neurotic, extroverted across situations (see any introductory personality or general psychology textbook; e.g., De Raad & Perugini, 2002; Gleitman et al., 2004) and there is little doubt that these general traits extend to emotional 'problems.' It is self-evident that employing an identical interpersonal approach to nearly all situations is generally a poor idea and there is some intuitive appeal to suggesting that part of EI is the ability to be *flexible* in solving an emotional dilemma. We normally associate EI with people who can deal with a variety of situations in different manners – good leaders adjust their leading style to accommodate the needs of their subordinates. The *awareness* mentioned above is important for determining what potential situations there are, but without *flexibility* in response to this perception, no real effect is discerned.

This conjecture may be construed as being too similar to some existing constructs, notably self-monitoring and the blanket term 'social skills,' but this is not the case. A close inspection of self-monitoring scale (Snyder, 1974) items reveals that this construct has much more to do with manipulation and façade than it does about emotions. The self-monitoring scale does not test the component abilities of EI, let alone the *awareness* and *flexibility* suggested here, thus it is unlikely that this conjecture is a simple re-packaging of self-monitoring. It is possible that this conjecture overlaps with 'social skills' but this seems unlikely. Frankly, the 'social skills' label is sufficiently broad to apply to nearly any interpersonal ability (EI, personality, perhaps even IQ, etc) and because it is broad and rarely defined in a precise manner, it is difficult to even know what such a claim means. It is not likely that there is a 'social skills' test which measures what this conjecture proposes, so it is difficult to argue that the *awareness* and *flexibility* conjecture is psychometrically redundant. Perhaps the claim is that this conjecture merely presents a single skill that is part of a larger set of 'social skills,' and therefore the conjecture is not worth investigating. But this is hardly a convincing argument against this conjecture. It would be less than compelling to argue that IQ a negligible concept simply because it falls under a broad umbrella term such as 'human abilities', and thus by analogy it is probably silly to argue that this conjecture fails simply because it might fall under an umbrella term like 'social skills.'

An example might be useful in clarifying this conjecture but because this conjecture is concerned with behaviour across situations or individuals, multiple situations must be represented. In the previous example of confrontations with a loved one, according to this conjecture a low-EI individual would be expected to be oblivious to his/her partner's emotional state or other subtle situational variations during these confrontations and consequently, would handle each such confrontation in exactly the same manner. It is possible that in addition to across-situation similarity, low-EI individuals would display across-individuals similarity: different low-EI individuals handle confrontation in a similar manner, despite the subtle differences between these confrontations (e.g., a caffeine-deprived partner or a happy, confident partner, etc.). In other words, low-EI is characterised by internal and perhaps external similarity. High-EI would be characterised by an increased awareness of the situational variability that characterises emotion-laden situations and consequently, greater internal variability. High-EI individuals would be less likely to handle similar situations in a similar manner. Rather, they would be more likely to note the subtle differences between (what would appear to low-EI individuals as similar) situations and behave variably in response to them. There is also likely to be little similarity between high-EI individuals in how they address such confrontations. In other words, according to this conjecture low-EI individuals are likely to be internally (same person, across situations) and externally (different people, same situation) similar and high-EI individuals are likely to be internally and externally dissimilar. This conjecture leads immediately to testable hypotheses.

3.4 Experimental Applications

This conjecture leads to a number of clear hypotheses which fall into at least two categories: test-based hypotheses and experiment-based, EI skill-related hypotheses. A number of these will be discussed and some will subsequently be tested in this thesis.

There are clear testing implications for the *awareness* and *flexibility* conjecture. If high-EI individuals see greater subtlety and variability in emotional situations and respond more flexibility in response to this variability, it is would be expected that high-EI individuals would respond more variably to EI test items than low-EI individuals. *Trait* EI test items (e.g., 'I am good at handling my emotions', 'People find it easy to confide in me') are simple statements that are written to represent and distil a variety of situations. Because of their greater *awareness*, high-EI individuals might read the item in numerous ways. In other words, the external variability proposed to characterise high-EI individuals would be visible on EI tests. On the other hand, low-EI individuals would be expected to view these items in a similar manner to one another due to their external similarity. In psychometric terms, these differences between sub-groups in item interpretation would be expected to lead to differences in factor structure and the explanatory power of each factor. This hypothesis will be explored in chapter 4. Another competing possibility is that high-EI individuals select more 'neutral' self-report responses on *trait* EI tests, owing to the fact that they are more likely than low-EI individuals to claim that their response 'depends on the situation.' It would also be worthwhile seeing whether or not a test designed to test EI as envisaged by this conjecture predicted any useful outcomes and this possibility is examined in chapter 8.

There are also numerous experimental implications, but given that these skills are numerous (especially in some *trait* EI theories) for brevity's sake, only affect regulation will be discussed here. One immediate hypothesis is that given an emotional situation with some degree of ambiguity, high-EI individuals would be expected to use affect regulation variably whilst low-EI individuals would be expected to regulate affect in a predictable manner due to the greater number of ways high-EI individuals 'read' a situation and respond *flexibly* in response to it. This hypothesis is tested in chapter 5. If this is the case, it is worthwhile to see if manipulating the clarity of the situation removes this ambiguity – if they are told precisely what to do, high-EI individuals are probably better at affect regulation. This possibility is examined in chapters 6 through 8. There are numerous other hypotheses which follow from this conjecture. In a free-response paradigm, it would be expected that high-EI individuals could list a greater number of situations in which affect regulation would be a 'good' or a 'bad' idea. The assertion that high-EI individuals 'read' situations better than low-EI individuals could be tested by asking high-EI individuals to list the ramifications and solutions they see in an emotional situation and determining whether or not their lists are longer and more detailed than low-EI individuals'.

3.5 Current EI Theory

Throughout this thesis, comparisons between 'current' or 'present' EI theory and the conjecture are made. Of course, given the amount of debate and controversy about this construct, it is probably inaccurate to say that there is simple *one* EI theory. If nothing else, there is at least *trait* and *ability* EI to contend with and it is not clear which this phrase refers to.

The predictions made in this thesis were designed so that they would follow from either *trait* or *ability* EI. The studies here deal with either affect regulation or psychometric variability and both types of EI theory would make similar predictions for these two outcomes. Specifically, both *trait* and *ability* EI theories would predict some sort of connection between EI test scores and reactivity to affect induction – either a positive association suggesting that high-EI individuals are more sensitive to affect or a negative association suggesting that high-EI individuals are better at regulating their affect – and both theories would probably predict that EI tests measure people on the same continuum. So the phrase 'current EI theory' is an oversimplification but it is used so that it is clear that the predictions based on 'current EI theory' were designed to fit well in both *trait* and *ability* EI theory.

That said, a pure *trait* theorist might argue that this thesis primarily deals with *ability* EI theory because some of the studies investigate a EI-related skill. Certainly, affect regulation is a skill which seems to fit better into *ability* EI, which postulates that EI is a collection of skills, than *trait* EI, which postulates that EI is a collection of behavioural tendencies and self-beliefs. Strictly speaking, *trait* EI does not demand that high-EI be equated with objective indices of affect regulation because this type of EI does not claim that EI is a set of skills. This is similar to the fact that personality theorists do not claim that, for example, extroverts actually *are* more socially apt. However, these same personality theorists most certainly do argue that extroverts believe that they are socially skilled and that these self-beliefs are often associated with real-life performance. A similar situation is actualised in these studies. *Trait* theory does not predict that high EI is equated with greater affect regulation, but it does predict that self-beliefs about one's tendency to regulate affect would have some sort of connection to objective indices of affect regulation. If this were not the case, the experiments described in Petrides and Furnham's (2003) would not make any sense.

3.5 Chapter Summary

In this chapter, a deficiency in current EI theory was noted and the *awareness* and *flexibility* conjecture was presented to address this deficiency. This conjecture complements, rather than opposes, current EI theory, and simply suggests that in addition to the EI skills already proposed, the ability to discern multiple solutions to emotional problems based on situational variations in seemingly similar dilemmas and the ability to be flexible in response to these different potential solutions. The experimental implications of this are clear and a number of hypotheses were discussed, including several that are tested in chapters 4-7.

Chapter 4: A factor analytic investigation of the *awareness* and *flexibility* conjecture

4.1 Introduction

As was previously mentioned, the conjecture immediately leads to a factor analytic hypothesis. The purpose of this study was to test this hypothesis using several large data sets from numerous projects.

4.1.1 The Logic of the Current Study

The *awareness* and *flexibility* conjecture suggests that EI is more than simply a collection of abilities which people differ in. Rather, EI is also the differences in how *aware* people are of which EI skill is most appropriate in any given situation and how *flexible* people are in employing their array of EI skills. Thus, a high-EI individual is one who uses different skills in different situations, not one who always employs one 'emotional technique' regardless of situation. In other words, high-EI individuals use their EI skills differently than low-EI individuals.

The logic behind this study and its relevance are explained below after certain assumptions are set out. It is a core assumption of classical test theory that tests measure high- and low-scoring individuals on the same dimension. In other words, no matter how high or low an individual's ability is, it is assumed that the test still measures the same ability. For example, it is assumed that a ruler measures tall and short people on the same dimension (i.e., height) and if this were not the case, we would probably be very sceptical about the utility of that ruler. It has not been determined whether or not EI tests meet this assumption but it seems that one of the tacit assumptions of current EI theory is that they do indeed display this 'factorial invariance'.

The conjecture makes a slightly different prediction because EI test items contain considerable ambiguity. For example, the item 'I can deal effectively with people' could represent numerous social and personal situations (e.g., at a party while feeling gregarious, at home while relaxing, etc) and it is left up to the test-taker to contextualise the item appropriately. This is no different from personality items.

Thus the logic for the present study works as follows:

(1) If the conjecture is true, then EI should be understood as greater *awareness* of the 'emotional approach' to take in any given situation and greater *flexibility* in response to it. In other words, given a diverse collection of situations, high-EI individuals would be expected to behave differently (more variably) from low-EI individuals in response to this collection of situations.

(2) EI tests are ambiguous, thus they (by definition) represent a multitude of possible 'situations.'

(3) Therefore, high-EI individuals would be expected to behave differently (more variably) from low-EI individuals in response to EI tests.

(3) can be thought of a 'differentiation' hypothesis and will serve as the null hypothesis for this study. The single positive prediction for this study is based on the 'factorial invariance' assumption above: that EI tests will measure high- and low-scorers on the same dimension. The conjecture-driven hypothesis is clearly the opposite of what would be expected based on current EI theories. Yet it is still of interest whether or not EI tests meet the factorial invariance assumption, even if the conjecture is not supported in this study. Thus the findings of this study bear on EI theory regardless of whether the *awareness* and *flexibility* conjecture makes accurate predictions.

4.1.3 Differentiation

Spearman (1927, cited in Deary et al., 1996) introduced psychometric differentiation in his 'law of diminishing returns' (LODR). Differentiation simply refers to a difference in behaviour between two groups. In the case of LODR, it refers to a specific difference between high- and low-IQ individuals: *g* would predict less variance in the IQ scores of high-IQ individuals than in low-IQ individuals. He expected that there would be differentiation by IQ score because high-IQ individuals use their intelligence in different manners – some intelligent people read Shakespeare and some arrange Schubert.

LODR has been equivocally supported since Spearman's initial suggestion. For example, Deary et al. (1996) found differentiation by IQ score in a large general population sample although it seems that the extent of the differentiation depended on the IQ sub-scale used to create sub-groups. Additionally, Deary et al. (1996) showed that by creating several sub-populations (i.e., more than simply a high- and low-scoring group) one could reveal more fine-grained differences in the explanatory power of *g*. Jensen (2003) found similar results but argued that only two sub-groups should be formed and it is acceptable to create them by using overall IQ scores. However, Hartmann and Teasdale (2004;2005) found that differentiation did *not* occur when they examined test scores from a very large sample of the Danish general population. Differentiation has been explained by using an economic analogy: perhaps people high in IQ 'spend' their intelligence in a variety of different manners (e.g., maths or verbal skills) and as such *g* explains less variance than amongst those low in IQ who are more 'bound' by their low IQ. In other words, high IQ leads to an increased role of preference, but low IQ individuals lack this freedom and their responses are better predicted by *g*.

Other EI-related constructs also seem to show differentiation of some kind. Personality tests appear to differentiate by IQ level (Austin, Deary, & Gibson, 1997; Austin, Hofer, Deary, & Eber, 2000; Harris, Vernon, & Jang 2005) with high-IQ individuals displaying more variance in personality item responses. This area of research has been less well-examined but these findings do appear fairly clear. This differentiation by IQ could occur because intelligent people detect more subtlety in personality items (similar to the *awareness* suggested) and consequently respond more variably to them (Austin et al., 1997).

Regardless of whether the economic metaphor or the 'subtlety' explanation work, or even if differentiation findings are robust, the methodology for differentiation studies has been clarified recently. Research in these areas also provides additional justification for examining differentiation in EI tests beyond the *awareness* and *flexibility* metric. The economic metaphor could easily be extend to include EI: high-EI means being capable of 'spending' one's EI to be a good affect regulator or being better at emotion perception in others. Or it could be that high-EI means being better at detecting the subtlety of EI test items. This study will proceed from the *awareness* and *flexibility* conjecture but these possibilities are explored in the discussion section.

4.1.4 Differentiation Methodology

The most detailed recent articles (Hartmann & Teasdale, 2004; 2005; Jensen, 2003) present a clear method to studying differentiation and these methods will be used for this study.

There are two primary questions that a differentiation study can be used to answer. Firstly, it can be determined whether or not the factor structures of the test scores of the two sub-groups are similar. That is, by using congruence coefficients (CC) and Cattell & Baggaley's (1960) salient similarity index, *s*, it is possible to see whether the item factor loadings of the factors in two sub-groups are similar. If the item factor loadings are dissimilar, then it can be stated with some certainty that the factor structures differ between the sub-groups.

Secondly, it is possible to see if the first unrotated principal axis factor explains more or less variance in two sub-groups. Harmann and Teasdale (2004) advocate the use of principal axis factoring (PAF) instead of principal components analysis (PCA) because PAF examines only the common variance of the correlation matrix, rather than both the common and unique (error) variance in the correlation matrix. Regardless, the *g* estimate extracted from this procedure is informative only if the factor structure of the sub-groups is nearly identical (Jensen, 2003 suggests $|CC| > .95$), otherwise one is effectively comparing apples and oranges.

Assuming the factor structures are nearly identical, the explanatory power of *g* in each group can be converted into an *F* ratio by simply dividing the *g* estimates (higher variance-explained value as the numerator). This *F* ratio can be tested for significance by using $F(n_1\text{-}1, n_2\text{-}1) = X$ (where n_1 is the *N* for the group expected to have more variance and n_2 the *N* for the group expected to have less). Degrees of freedom terms are large because each *g* value is a variance estimate *within* groups, not between groups.

Some other precautions must be taken in order to ensure that the *g* estimates extracted from factor analysis. The cut-off point (e.g., IQ score of 95) on the test used to create sub-groups must ensure nearly equal standard deviations on this outcome measure (e.g., IQ standard deviations in both sub-groups of 15.5 and 16.1). Secondly, a reliability estimate must be taken for the item responses in both sub-groups to ensure that any differences in the variance explained by *g* are not simply due to overall differences in variability between groups.

Essentially, the procedure is as simple as dividing the sample into two subgroups based on a cut-off score that ensures equal standard deviations in both groups, factor analysing the test results of the sub-groups with PAF, calculating CC's and *s*, then calculating and comparing the variance explained by *g*. This procedure will be used for the EI test results examined for this study.

4.1.5 *S*, the Salient Similarity Index

Cattell and Baggelley (1960) introduced this statistic, *s*, as a non-parametric alternative to CCs. In order to calculate either a CC or *s*, a factor analysis must be performed and the item factor loadings for any relevant factors must be examined.

The method for calculating this is explained in detail elsewhere (Zack, Toneatto, & Streiner, 1998) but the easiest approach is to count the number of irrelevant (i.e., either or both factor loadings are $|r| < .1$), concordant (i.e., *both* loadings are $|r| > .1$

and have the same sign – e.g., .40 and .35), and discordant (i.e., both loadings are $|r| > .1$ and have *different* signs – e.g., .40 and -.35) loadings. These loadings then allow the *s* value to be calculated for each factor in each sample as follows:

$$s = \frac{c - d}{[c + d] + .5i}$$

Where *c* is the number of concordant loadings, *d* is the number of discordant loadings, and *i* is the number of irrelevant loadings.

This index ranges from -1 to 1 and is interpreted in roughly the same manner as a CC.

4.1.6 Hypotheses

The *awareness* and *flexibility* conjecture make several immediate predictions. The first is that because EI tests are made of items that are ambiguous, the factor structure of high-EI individuals' responses will be dissimilar to that of low-EI individuals' responses. Specifically, based on this conjecture it is predicted that the CC and *s* values for the factor structure of the two subgroups will be considerably less than identical (i.e., $|CC| < .8$). Additionally, it could be predicted that although the factor structure of the subgroups' responses will be highly similar, there will be a difference in how much variance the first principal axis factor (known hereafter as emotional *g*) will explain in the two subgroups. This hypothesis depends upon the confirmation of the first and will be tested in turn. Current EI theory would predict that the factor structures will be highly similar and because this is a positive prediction, it will serve as the 'experimental' hypothesis while the conjecture-driven hypothesis will serve as the 'null' hypothesis.

For the sake of clarity and because the same method was used for all samples, only one method section will be presented here.

4.2 Method

4.2.1 Participants

A total of five data sets which were collected at different times in either Scotland or Canada. With the exception of Samples 1 and 2, which completed the same questionnaire, all completed different measures. Details on the measures are below.

Sample 1 was 196 University of Edinburgh students. Twenty six did not provide gender information, 51 were male and 121 were female. The participants ranged in age from 18 to 46, though the vast majority were aged between 18 and 21. These participants completed the 41-item variant of the EIS.

Sample 2 was comprised of 623 Canadian (445) and Scottish (178) undergraduates of which 431 were female and 187 were male and five did not disclose their gender and whose ages raged from 16 to 81 ($M = 28.5$, $SD = 14.8$). These participants are identical to the sample used in a recent study (Austin, Saklofske, & Egan, 2005) and completed the same EIS variant as Sample 1.

Sample 3 was comprised of 500 participants of either Canadian (340) or Scottish (160) nationality for a recent study (Austin & Saklofske, 2005). There were 374 females and 126 males in this sample and ages ranged from 17 to 59 ($M = 24.2$, $SD = 6.6$). These participants took the un-adapted 33-item EIS.

Sample 4 was made of 537 participants of Canadian nationality who were sampled as part of the same research programme as Sample 2 (Austin, Saklofske, & Egan, 2005). These are different participants from Sample 2. 377 of the participants were female, 156 were male, and 4 did not provide gender information. Ages in this sample ranged from 18 to 58 ($M = 24.3$, $SD = 6.8$). These participants completed the short-form Bar-On EQ-i test.

Sample 5 was of 99 Edinburgh University students. Seventy were female and 29 were male. The mean age was 22.4 years (*SD* = 8.0). These participants completed the MSCEIT v2.0 (Mayer et al, 2002) as part of a study that is currently in preparation (Austin, submitted) This ref needs to be changed, it is Study 1 from Farrelly & Austin (in press).

4.2.2 Measures

33-item EIS. This measure, (EIS; Schutte et al, 1998), is a 33-item self-report (5-point Likert) measure which assesses the EI framework proposed by Salovey & Mayer (1989): overall EI or the Optimism/Mood Regulation, Utilisation of Emotions or Appraisal of Emotions subfactors. Some of the items are reverse keyed and its internal reliability here is high ($\alpha = .89$).

41-item EIS. This variant of the EIS, developed by Austin et al. (2004) has slightly more items and more reverse-keyed items. This test has 41 items, of which roughly half are reverse-keyed. It is answered in the same self-report manner as the 44-item EIS. The total scale internal consistency here was acceptable ($\alpha = .85$).

EQ-i Short Form. (Bar-On, 2002) This trait EI test measures total EI and five EI sub-factors: Interpersonal EI, Intrapersonal EI, Adaptability, Stress Management, and General Mood. There are also several items that check for 'positive impression' bias which were omitted from the analyses here as they are irrelevant to the hypotheses. This test has 51 items in total and had high internal consistency in this sample ($\alpha = .90$).

MSCEIT v2.0. This test is owned and distributed by Multi Health Systems and is reported by its authors to have high (i.e., $\alpha > .8$) internal reliability (Mayer, Salovey, & Caruso, 2002). MHS does not furnish researchers with item data, and as such analyses have been performed on the sub-scale scores which are returned by MHS: faces, pictures, sensations, facilitations, blends, changes, emotion management, and emotion regulation.

4.2.3 Procedure

Participants completed the tests at different places and times. Sample 5 completed the online MSCEIT 2.0 and the remaining samples completed pen-and-paper EI tests. All were given instructions on how to complete the questionnaires and were informed of the purpose of the study in debriefing. Additional detail is available in the above articles.

To test the hypothesis, high and low groups were created. Because the sub-scales of EI tests are contentious (e.g., Petrides and Furnham, 2000, Davies et al., 1998, Ciarrochi et al., 2000), groups were separated by an overall EI score that ensured equal standard deviations in both sub-groups (after Jensen, 2003).

4.3 Results

Descriptive statistics show that these samples display the well-established gender difference. Table 4.1 shows a breakdown of the samples, the cut-off score that was used to divide them into high- and low-scoring groups, and also presents a significance test for the male-female EI difference.

Table 4.1 Cut-off scores, descriptive statistics, and male-female differences for the five samples

Sample	Test	EI >, <= Score	Standard Deviation		N		F-M T-test
			Low EI	High EI	Low EI	High EI	sig
1	EIS-44	145	7.11	7.12	84	112	*n.s.*
2	EIS-44	144	10.39	9.99	146	477	**0.001**
3	EIS-33	119	10.96	7.68*	201	299	**0.001**
4	EQ-i	31	2.33	2.24	141	396	**0.001**
5	MSCEIT	95	5.34	5.29	44	55	n/a

Note: EIS-44 refers to the 44-item EIS variant, EIS-33 refers to the original 33 item EIS, EQ-i refers to the EQ-i short form, MSCEIT refers to the MSCEIT v2.0.

EI>, <= Score refers to the cut-off score from which greater than or less than or equal-to groups were made on that test. F-M T-test sig refers to the significance of the t-test carried out between males and females; females scored higher than males in all tests. This test was not carried out for the MSCEIT data due to small sample size.

Factor structures differed between EI tests. To find appropriate solutions for these data, principal axis factoring with oblique rotation was used (after Austin et al., 2004) because EI has historically been argued to be a hierarchical construct. There was limited similarity in factor solutions between tests. The scree plots indicated a three-factor solution for samples 1 and 3, and a two-factor solution for samples 2 and 3. The MSCEIT data favoured a three-factor solution as no additional factors had initial eigenvalues greater than 1. In all samples, additional factors explained less than 5% of the variance and in samples 1 through 4, the first factor was the only factor to explain more than 10%. Five and even six-factor solutions could be defended on the grounds that additional items load uniquely on them. However, a factor that only accounts for one variable is hardly important, especially in light of the fact that some items did not load meaningfully on *any* factors. The solutions here for samples 1, 2, and 3 are similar to previous work (e.g., Austin et al., 2004; Petrides & Furnham, 2000) and the solution for sample 4 is novel but the solution for sample 5 differs from the four-factor structure previously found (Parker et al., 2005), possibly
because only sub-scale data could be analysed.

As Table 4.1 shows, the sub-groups were split such that sub-groups had nearly identical standard deviations. All standard deviation values were converted to an *F*-ratio (high group standard deviation)^2/(low group standard deviation)^2 and tested using a similar procedure explained above. All differences in standard deviation between the two sub-groups were nonsignificant with the exception of Sample 3. The differences in standard deviation in Sample 3 should be taken into account during discussion of differences in the explanatory power of emotional *g*.

To test the first hypothesis, two indices of factor similarity were calculated. The first is a CC which is calculated by considering the sub-groups' item factor loadings as two separate continuous variables and then correlating them with one another. This process can be repeated for as many factors as is desired. The second index is the *s*, discussed above.

It was hypothesised that indices of factor similarity would show that the factor structure of the two tests would be markedly different. Jensen (2003) suggests that for an IQ test, any CC below .9 indicates that the factor structure is notably dissimilar. Because EI tests are less robust than IQ tests in a variety of ways (including factor clarity), a less stringent criterion (perhaps $|CC|<.8$) could be used to determine whether the factor structures are similar or not. The first three factors are examined for all five samples for two key reasons. First, based on scree plots, eigenvalues, and estimates of variance explained, only a three-factor solution accurately described all of the responses. Secondly, additional factors explain approximately the same quantity of variance (i.e., factors 4,5, and 6 each add 3% of the variance) so inclusion of additional factors would be somewhat arbitrary.

Table 4.2 displays the indices that were calculated as well as a significance test of the *F* ratio of the variance explained by emotional *g*. What is immediately apparent is that none of the comparisons met the criteria set out here or in Jensen (2003) and that most factors are notably dissimilar (mean $|CC| = .40$, mean $|s| = .34$).

It is therefore not justified to examine the explanatory power of the first unrotated principal factor. The results indicate that the dimensions which item responses load differ between the two sub-groups and as such any comparison of these factors would be analogous to comparing apples and oranges. Moreover, the low CC and *s* values already indicate that the factor structures of the two sub-groups are different. Calculating and comparing the variance explained by the first unrotated principal factor would add little to our understanding of factor structure differences and is not, strictly speaking, acceptable given the low CC and *s* values.

Table 4.2 Congruence coefficients, *s* index values, and differences in variance explained

Sample	Test	Factor 1		Factor 2		Factor 3		Factor 4	
		CC	*s*	CC	*s*	CC	*s*	CC	*s*
1	EIS-44	-.09	-.19	.27	.12	-.10	.08	-.30	.08
2	EIS-44	-.37	-.25	.19	.19	.23	.28	n/a	n/a
3	EIS-33	-.15	.32	.70*	.40	.42	.20	n/a	n/a
4	EQ-i	-.20	-.12	-.07	.11	.82*	.25	-.45	.03
5	MSCEIT§	.31	.57	.75*	.15	.29	.33	n/a	n/a

Note: PAF1,2,3 refer to the first, second, and third unrotated principal axis factor

CC refers to the strength of the congruence coefficient calculated between high- and low-EI groups. *s* refers to the strength of the salient similarity index.

* acceptably high to consider as an identical factor loading structure (Jensen,2003)

§ MSCEIT analyses were run on sub-scale data

Because women score higher than men on EI tests in these samples, the high- and low-scoring groups in these samples were confounded with gender. Men and women seem to show different factor structures as well as their test scores seem to show differences in factor structure (average |CC| = .43), so it would appear that gender differences and EI score differences are intimately linked. With this caveat in mind, the results seem to indicate that the null hypothesis (i.e., the conjecture-driven hypothesis) could not be rejected.

4.4 Discussion

Only one of the indices of factor similarity met the criteria for factor similarity. These findings do not allow the rejection of the null hypothesis that there would be no correlation between factor loadings. However, this must be tempered by the finding that the responses of gender sub-groups had different factor structures and the fact that the high-scoring group was made mostly of women while the low-scoring group was made mostly of men. In other words, gender was a confound for the findings here, it is not clear whether the factor structure differences are due to differences in EI or differences in gender. This problem besets much of EI research as the gender difference is robust. A replication using high- and low-scoring groups with equal gender composition would be fruitful as would a more detailed exploration of why this gender difference occurs and what importance it has for EI theory.

Alternative explanations for these findings exist. Because EI scores correlate with personality and IQ, it is possible that any findings here could have simply been due to personality or IQ differences. Moreover, it is possible that a mechanism similar to those proposed to explain IQ or personality differentiation (i.e., the economic or 'subtlety-detection' explanations) operate in EI. It is difficult to determine if any of these should be endorsed in favour of the *awareness* and *flexibility* conjecture, not least because this is, to my knowledge, the first study of factor structure differences in EI tests. However, these explanations seem to have less currency simply due to reasons of parsimony and relevance.

It is possible that the results of this study are due not to factor variability but insensitivity of CCs or *s*. The strength of these correlations might have been low due to low reliabilities, poor methodology, or other variables and might have nothing to do with factor structure. These correlations are relatively unsophisticated compared to techniques such as confirmatory factor analysis and are less than ideally suited to this study. Still, the hypotheses are simple so a simple analysis is probably acceptable even if more sophisticated analyses might be more revealing. Current EI theory would predict that EI tests measure high- and low-scoring individuals on the same dimensions and these simple correlational procedures fail to confirm this in these data. It could be argued that because the CC and *s* values are so weak, it is unnecessary to look any further with more sophisticated techniques. However, techniques such as CFA or structural equation modelling could be used to determine whether or not the weak CC and *s* values are due to extraneous variables, so there is certainly room for their use in future studies. These correlations are too weak to indicate factorial invariance but it is certainly possible that these correlations were weak due to the presence or operation of other variables.

The economic and 'subtlety detection' mechanisms proposed to account for IQ and personality differentiation, respectively, have been proposed to explain difference is the explanatory power of the first unrotated factor, *not* differences in item factor loadings. In order to use these mechanisms to explain item level differences, an additional (or different) line of reasoning would be necessary. The same applies to the claim that any differences in EI test factor structure are due simply to differences in IQ or personality – additional reasoning would be necessary to explain how these differences would cause both differences in EI scores *and* differences in factor structure between EI groups. In contrast, the conjecture presented here predicts all of these findings without any additional explication. Obviously a great deal of experimental research is necessary before a firm conclusion is drawn but if nothing else this conjecture receives tentative psychometric support in the sense that it was not possible to reject the null hypothesis which followed from the conjecture.

A more important issue is whether or not the failure to reject the null hypothesis (i.e., factorial variability, no correlation between factor loadings) should be equated with retention of the null. That is, although it was not possible to reject the conjecture-driven hypothesis (a null prediction of factorial variance) in favour of current EI theory, it is not necessarily the case that the conjecture-driven hypothesis was accurate. For the reasons mentioned above, the failure to reject the null hypothesis could be due to other variables and the simple methods used in this study are powerless to reveal whether the null finding was due to factorial differences or any other variable. So although these results seem to contraindicate current EI theory, this is not a final conclusion.

Another issue with this study was the division of sub-groups. The high- and low-scoring groups were created based on their overall EI test scores. However, participants were divided on the same tests which are later showed to lack factorial variability and thus the method is slightly self-referential. It hardly seems appropriate to argue that the tests are flawed but are also acceptable for sub-group creation. This criticism is mitigated by several comments. Firstly, it would have been necessary to know the results of the study in order to know that the tests are factorially variable, thus the use of a different division criterion would be, by definition, *post-hoc*. Secondly, if sub-group division was not allowed to be based on EI test scores, then it is hard to say what exactly it should have been based on. Thirdly, other studies create sub-groups based on scores on tests which show evidence of differentiation (e.g., Jensen, 2003). Perhaps an EI test not analysed in the study could be used for division, but this would force a question about why this 'selection EI test' was not analysed. Also, the conclusion that 'EI tests are flawed' is a rather harsh conclusion in light of the possibilities discussed above, so if one wishes to argue that sub-group division based on flawed EI tests is questionable, then one must tacitly assume that these results indict EI testing in general. In a way, this criticism of the division methods used in this study leads to the very conclusion that the criticism is designed to defend against: that current EI tests are flawed. It would seem that there is certainly something dubious about dividing by one criterion and then later arguing that that criterion is flawed, and although there don't seem to be many alternatives, it could be that using peer ratings would address this issue.

That said, a major issue with this study is that it is not clear what is driving the factorial differences observed here. These differences might have been observed if participants were divided according to their personality scores and as was already noted, gender might have played a role. This could be problematic because it introduces an ambiguity as to why the conjecture does or does not make accurate predictions in these studies.

A replication which statistically or otherwise controlled for intelligence and personality when evaluating EI test scores would be ideal. More complex modelling packages could be used for this purpose and would be useful in ascertaining whether or not any of the prior alternative suggestions should be endorsed in favour of the conjecture presented here.

Additionally, sampling could have been improved in this study. The samples used were not as large as have been used is recent IQ differentiation studies (e.g., Hartmann & Teasdale, 2004). A larger, more representative sample would help rule out gender differences by allowing the analysis of high- and low-scoring groups with equal numbers of males and females. Such a sample would also circumvent the non-representative nature of university students, however EI was roughly normally distributed in the samples (some negative skew occurred) and because IQ and *trait* or *ability* EI tests are at best moderately correlated with one another, restriction of the IQ range should have no important repercussions. Moreover, restriction of sample range does not *necessarily* result in an underestimated population effect size. An similar investigation of test responses of clinical (e.g., alexithymics) groups would shed additional light on the viability of the conjecture presented here.

Perhaps the most important flaw in this study is, as was previously discussed, the groups were split on the same measure which was later shown to be problematic. The most obvious adjustment would be to replicate this study but factor analyse a second variable rather than the EI test scores. If a different, external variable (e.g., coping behaviour) was analysed in a similar manner to determine if high-EI individuals are more variable, it would make it more sensible to divide according to EI test scores and would also avoid a situation where a single variable performs more than one purpose.

Additionally, experimentation will be necessary to determine whether or not this conjecture functions outside of the rarefied environment of psychometric testing and as such the following three chapters are all experimental. Still, there is evidence here for the *awareness* and *flexibility* conjecture in a number of samples which were collected in different countries at different times. And even if this evidence is tentative, it is still interesting because it stands in direct opposition to present EI theory which would assume that EI tests measure identical dimensions for all test-takers.

4.5 Summary

In this study, participants in several samples completed EI tests. High- and low-scoring sub-groups were created based on their EI test scores and their responses to EI tests were then factor analysed and the factor structures of their responses were examined in detail. It was predicted according to current EI theory that the factor structures of the responses of the two groups would be highly similar, that the test would be factorially invariant. The opposite prediction was made according to the conjecture. It was observed that the factor structures of responses from the high- and low-scoring sub-groups were markedly different such that item factor loadings were uncorrelated in the two sub-groups. Thus it was not possible to confirm the prediction made according to current EI theory.

Chapter 5: Un-prompted Affect Regulation in Response to Mood Manipulation

5.1 Introduction

The purpose of this study was to determine if individuals with higher EI were less reactive to affect inductions than individuals low in EI. An introduction to SCR and important affect terms such as valence and arousal is presented, along with hypotheses that follow from the *awareness* and *flexibility* conjecture and current EI theory.

5.1. Operationalising and Inducing Affect

As was explained in chapter 1, this thesis is chiefly concerned with the physiological aspects of affect. The specific concern is used throughout this thesis though it should be noted that this definition does not preclude other definitions of affect. Indeed, in chapters 7 and 8, a self-report index of affect is used alongside the physiological index.

There are numerous ways to induce affect. Participants can be shown film clips, pictures, music, their own pre-written affective memories, or they can even be asked to adopt a facial expression. These techniques have been compared to one another along a number of dimensions (Martin, 1990) such as the strength of affective changes, risk of demand characteristics, and the robustness of the procedures across studies. Some of the more obscure methods of inducing mood (e.g., computer programmes, listening to music) are often the least susceptible to demand characteristic biases but these techniques tend to result in smaller affective changes. In a meta-analytical study, Martin (1990) advocated the use of film clips as mood induction stimuli although she also noted that film stimuli suffer from demand characteristic bias because people who watch films with affective material generally know what sort of affect 'to expect.' That is, it is usually very clear what affective response a film has been selected to elicit, so it is possible that participants alter their reactions in response to this information. Because demand characteristic bias is probably easier to correct (by using e.g., double-blind designs, minimizing experimenter interaction, etc) than a weak experimental effect, Martin (1990) argued that demand characteristic issues are only of secondary importance for any decision on induction technique. In a similar meta-analysis, Westermann, Spies, Stahl, and Hesse (1996) suggested that film stimuli are superior to other affect induction techniques in terms of effect size.

There could be numerous reasons why film stimuli are so effective at inducing affect but these possibilities are only mentioned briefly here because this thesis is only concerned with induction, not the reasons why a particular technique is effective. It is possible that they are more effective because films stimulate more sensory modes than most other induction techniques. Films contain both auditory and visual content whereas other techniques are mono-modal (e.g., pictures and music) or take place solely in the mind of participant. Perhaps the additional sensory stimulation leads to additional affective stimulation. Regardless of the reasons, it is clear both scientifically and anecdotally that films are effective at inducing affect – were this not the case, it is doubtful that Hollywood films would be nearly as memorable. Because of these reasons, film stimuli are used to induce anxiety (negative affect) and happiness (positive affect) in this study.

5.1.2 Valence and Arousal

Before the details of electrodermal activity (EDA) or skin conductance response (SCR) are discussed, two important facets of affect must be discussed: valence and arousal. These two dimensions represent the bare minimum of what is required to describe an emotional response. That is, most theorists would acknowledge that any affective reaction must be described (at least) in terms of its positivity/negativity and how strong/weak the affective response was. Bradley (2000) explains how this two-factor distinction can be used to understand numerous affective responses, including reactions to IAPS pictures (which are used in this study). For example, pictures of babies are rated as equally positive, but less arousing than pictures of money and pictures of cemeteries are rated as more negative but equally arousing as pictures of jails. Of course, higher-level aspects of affect add more dimensions to this two-factor space, but even the most reductive, physiological account of affect requires some allusion to these two continuums. Bradley (2000) also suggests that some physiological indices are better at measuring one factor than the other – this property of physiological measurement of affect will be returned to below.

5.1.3 EDA and SCR/SCL

As Dawson et al. (2000) explain, EDA is a blanket term that refers to a number of processes in the integumentary, central nervous, sympathetic autonomic nervous, and peripheral nervous systems. Simply put, the skin resists an applied electrical current and the differences in this resistance between different situations (e.g., at rest or under stress) are a psychologically meaningful index of arousal levels. The skin's electrical resistance or its inverse, conductance, are measured over a period of time to observe a 'waveform' of EDA. Chief amongst the models suggested to explain how the peripheral mechanisms (e.g., sweat glands, sweat ducts, dermis, keratinised skin cells, etc) in the skin function to create electrical activity is the two-effector model which posits that the electrical activity is the result of a) the filling of the sweat ducts; and b) the activity of some selective membrane in the epidermis (Dawson, 2000). However, this model has been criticised because it postulates a membrane which, to be blunt, may not actually exist and has yet to be located (Edelber, 1993). There is also evidence from lesion studies in animals that EDA is caused by activity in the central nervous system, specifically the limbic system and amygdala (Boucsein, 1992), although these CNS pathways are numerous and complex. It was initially thought that EDA was controlled by both the parasympathetic and sympathetic branches of the autonomic nervous system, but it is now generally accepted that only the sympathetic branch controls EDA (Dawson, 2000). The immediate implication of this finding is that, unlike other physiological measures (e.g., heart rate, zygomatic muscle activity, etc), EDA is *solely* a function of the body's preparatory, 'flight-or-fight' response. As such, EDA is a measure of the body's state anxiety – increases in EDA suggest increases in anxiety and decreases in EDA suggest decreases in anxiety. These changes in anxiety/arousal are used as indicators of affect for this study and the remainder of this thesis.

There are two major approaches to the quantification of EDA. The first approach measures the tonic skin conductance level (SCL) over an extended period of time. The second measures the phasic skin conductance response (SCR) over a brief period of time, usually 5-10 seconds. In both cases, some kind of stimulus is presented to the participant, all that differs is the duration of the stimulus and the duration of EDA measurement. In the present study, a participant's reaction to a stimulus with a fairly long duration (one minute) is of interest, so SCL must be used. SCR is normally used for stimuli such as affective pictures and will be used in chapters 6 through 8.

SCL has strengths and weaknesses. SCL allows researchers to investigate the affective state of a participant over a longer period than SCR – it is possible to see not only individual, momentary reactions to stimuli but also longer-term trends in the participant's affective state. The cost of this increased sensitivity to longer-term trends is a loss of experimental control. With a longer stimulus it is not possible to ensure that every moment of the stimulus actually contains the affective material that is desired by the researchers. In the case of film stimuli, the dynamic nature of the medium itself prevents complete certainty that the *entire* stimulus is violent, humorous, etc. This lack of control has prompted many researchers to use SCR to the exclusion of SCL, but this thesis contains studies which use both methods. In this study, an increase in anxiety is defined as an increase in SCL throughout the affect manipulation and a decrease in anxiety is defined as a decrease in SCL. Because affect regulation has been previously defined as the ability to return one's affect to some individual baseline, in terms of SCL, a greater increase in SCL is indicative of *less* affect regulation and a smaller increase in SCL is indicate of *more* affect regulation.

It should be noted that EDA is a measure which has been used in innumerable studies of affective response. As was previously explained, it is far more sensitive to general affective arousal than it is to discrete types of emotion. This is the case for most physiological indicators – they are useful as gross indicators of affective activity but lack resolution for complex emotions. This does not in any way suggest that they are inappropriate dependant variables. On the contrary, the physiological aspects are affective response and emotion are virtually inseparable from the cognitive aspects. This thesis has been designed to measure general positive and negative affective response and as such EDA is a sensitive, reliable, and robust measure to use for these studies.

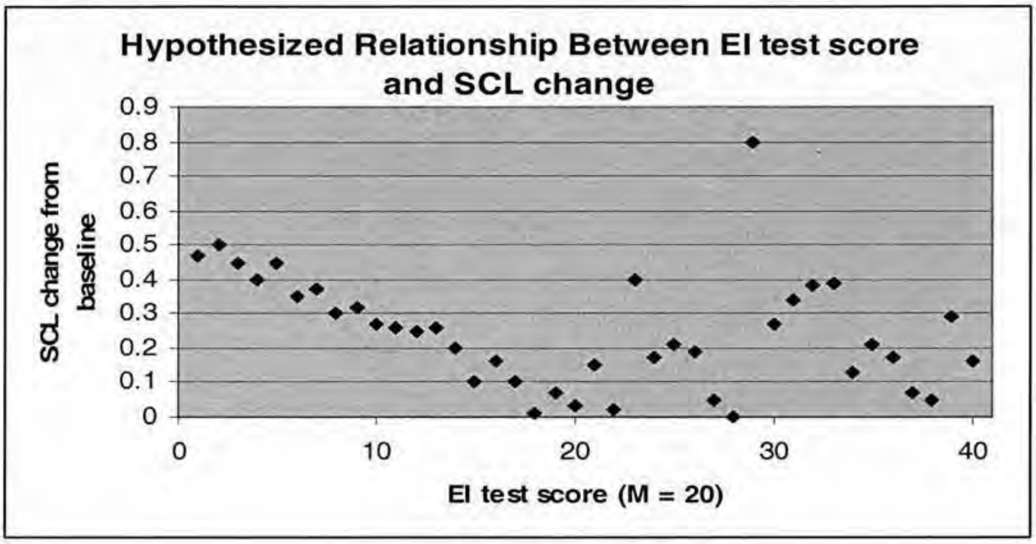
5.1.4 Logic of the Study

As was noted previously, affect regulation is measured indirectly in this thesis. Specifically, this thesis examines the link between EI test scores and affective reactivity to affect induction with the assumption that lesser reactivity is indicative of enhanced operation of affect regulation.

Prior studies which have examined the connection between EI test scores and affect regulation (Ciarrochi et al., 2000; Petrides & Furnham, 2003) predicted different outcomes based on different interpretations of current EI theory. Ciarrochi et al. (2000) failed to reject the null hypothesis in some comparisons and in others found that high-EI individuals are less influenced by affect manipulations when they perform a cognitive task and Petrides and Furnham (2003) found that high-EI individuals are more affected by affect manipulations when they simply report their affect. This may have something to do with methodological differences – if participants know that they will be required to perform a cognitive task they may employ their affect regulation abilities to a greater extent than if they simply report their affect. The differences in results between these two studies may also be due to the type of EI test used. It was acknowledged above that the logic in both previous studies of EI and affect regulation is equally compelling, and in spite of the methodological similarities between this study and Petrides and Furnham's (2003) study, for the sake of continuity, based on current EI theory it is predicted that higher EI test scores will be associated with lesser reactivity to the affect induction.

The conjecture makes a different prediction. Under this conceptualisation of EI, higher EI is not defined as simply more developed emotional skills, it is also a difference in how and when those skills are used. High-EI individuals are claimed to be more *aware* of the emotion content or implications of a situation and more *flexible* in response to this content. Low-EI individuals are claimed to be less *aware* and less *flexible*. In other words, the way in which high-EI individuals employ their emotional skills depends more on the situation than it does for low-EI individuals – given a somewhat ambiguous situation, high-EI individuals will behave more variably than low-EI individuals in response to greater *awareness* of the numerous ways that a situation could be approached. This was supported in chapter 4 which showed that high-EI individuals display a different factor structure in their EI test (i.e., somewhat ambiguous situation) responses. In terms of affect regulation, it would be expected that, given the same somewhat ambiguous situation, high-EI individuals would regulate more variably than low-EI individuals, because they 'read' a greater number of interpretations of this same situation and behave in different ways in response to it. In other words, due to greater *awareness* and *flexibility,* in a somewhat ambiguous situation (e.g., an affect manipulation not linked to any sort of cognitive task), high-EI individuals would be expected to *not* regulate their affect in any predictable manner. Some of the high-EI participants will view the films and decide to become immersed in the stimuli, and some will decide to resist the affective content. The low-EI individuals would be expected to regulate their affect in line with their EI level because, according to this conjecture, they would not have the same kind of immerse/resist *flexibility.* Briefly put, in the positive film condition, a positive association between EI score and arousal change from baseline is predicted for low-EI individuals. In the negative film condition, a negative association between EI score and arousal change from baseline is predicted for low-EI individuals. That is, amongst relatively low-EI individuals, it is expected that greater EI will be related to increased ability to maintain a baseline level of affect. No correlation between EI score and arousal change is predicted for high-EI individuals in any condition. Figure 5.1 illustrates what is predicted according to the conjecture.

Figure 5.1: An illustration of what is predicted in this study according to the conjecture



5.2 Method

5.2.1 Design

Trait EI was measured and subsequently Skin Conductance Level (SCL) was measured while participants rested (baseline) and watched both an anxiety-inducing (negative) and a happiness-inducing (positive) film clip in a correlational design. The order of the film clips was randomised to prevent order effects.

5.2.2 Participants

Forty-one participants (twenty nine female and twelve male) took part in this study. All were recruited using opportunity sampling and an advertisement placed on the University of Edinburgh Careers Service website. The questionnaire responses of these participants were also used in chapter 4 (sample 1). All were postgraduate or undergraduate students at the University of Edinburgh and varied in age from 18 to 41, with a majority aged between 18 and 21. One participant was excluded because her responses were very extreme (roughly four standard deviations from the mean) and were unduly influencing the overall results.They were paid £2.50 for their time.

5.2.3 Measures

The Austin et al. (2004) variant of the EIS was used in this study to test *trait* EI. It is a 41-item variant of the original 33-item EIS by Schutte et al. (1998) which contains a greater number of items of which more are reverse-keyed. Like the EIS, this is a self-report questionnaire on which test-takers respond to a series of statements with a 5-point Likert scale anchored at 'strongly agree' and 'strongly disagree'. The sample in this study was too small for an accurate estimate of reliability of the Austin et al. (2004) variant, but the responses of the larger sample (chapter 4) were highly reliable.

5.2.4 Apparatus

A computer program was designed by Dr. Paul Stevens of the Koestler Parapsychology Unit. This program presented all stimuli, measured SCL, and performed all calculations and transformations on the raw data. A touch-sensitive LCD screen was used as an interface for participants to advance through different stages of the experimental session.

Standard Skin Conductance measurement equipment was used to measure EDA throughout the experiment. 8mm sintered Ag-AgCl round cup electrodes filled with a pH balanced aqueous NaCl gel were affixed with double-sided adhesive collars. These electrodes were connected via a pre-amplifier to a model SC5-SA (Psylab/Contact Precision Instruments, London, UK) with 24-bit A/D conversion. Raw measurement was in microsiemens (muS) at approximately 40Hz and readings were z-transformed. All electrodes were place on the middle phalanx of the first and second digit of the participant's non-dominant hand to ensure that good contact was not prevented by calluses.

Two one-minute film clips were presented to all participants. The anxiety-inducing film (negative affect) was a scene from a recent movie in which a character vividly imagines a nuclear explosion destroying a city, leaving only the skeletal remains of the children she watches. The happiness-inducing film (positive affect) was a clip of polar bear cubs playing and rolling around in the snow. These clips were piloted before testing and all pilot participants reported feeling the affect that was expected.

Physiological data were z-transformed. This process is carried out to normalise and to allow for individual comparisons. Because there are massive individual differences in Skin Conductance – in terms of mean values and variance – it is not appropriate to simply calculate group means or examine correlations with another measure, especially in a between-subjects design. It is essential that these differences be controlled for and z-transformation suits this purpose. Instead of raw data, z-transformation essentially converts each raw data point into a z-score. Z-transformation was used instead of log or ln-transformation because the latter procedures do not allow individual comparisons.

The formula for this procedure takes the form of $Z_N = (X_N - X_M) / \sigma_N$ (where Z_N is the standardised score for each participant's raw data point, X_N is the raw data point, X_M is a participant's overall (including resting and experimental periods) mean level of arousal and σ_N is each participant's overall standard deviation of arousal level). By this process, each data point is expressed in terms of a participant's unique SD and mean. A participant's baseline mean and SD is calculated through the entire session, including experimental periods, so that the baseline is not artificially low due to being calculated from a resting period only. It is a well known transformation (Boucsein, 1992) that is also used in parapsychology research (e.g. Stevens, 2001) where sensitive physiological methods are commonly employed and individual comparisons are necessary.

5.2.5 Procedure

Participants were instructed to sit down and relax in front of the experimental monitor with the stipulation that they not cross their legs in order to prevent artefacts in the results. The minor change in blood flow that is caused by crossing one's legs creates a small difference in SCL, as does any body movement. Practical steps such as this one, as well as a 5-minute relaxation phase before mood manipulation to allow time for an accurate baseline and for the electrode gel to set into the skin properly are standard procedures.

While they were being connected to the EDA electrodes participants were informed about how the electrodes function and were assured that they could end the study at any time for any reason. After informed consent was obtained, each participant was attached to the electrodes and proper connection/reading was ensured.

The experimental session consisted of four periods. The first five minutes were a relaxation period. After this initial period, one of the two film clips was shown at random in a double-blind paradigm. Participants were instructed before each film clip to allow themselves to become immersed in the stimuli. Subsequent to this, each participant was instructed to relax for an additional two minutes, and then the remaining film clip was shown.

Following the experimental session, each participant was debriefed and compensated for his/her time.

5.3 Results

Table 5.1 displays the means and standard deviations for the arousal measures calculated as well as for EI test scores for this sample for reference. The data were non-normal even after z-transformation (Shapiro-Wilk testing confirmed this). This is expected of EDA measures and curtails methodological options to non-parametric methods such as Spearman's Rho. The descriptive statistics for the SCL measures are for reference only as overall group differences in SCL change magnitude are of no real meaning in this study. It is individual changes from baseline (i.e., relative differences) and how they relate to EI test scores that is of interest.

Table 5.1 Descriptive statistics for EI and SCL

Variable	M	SD
EI	25.72	14.86
Positive SCL	.041	.251
Negative SCL	-.043	.249

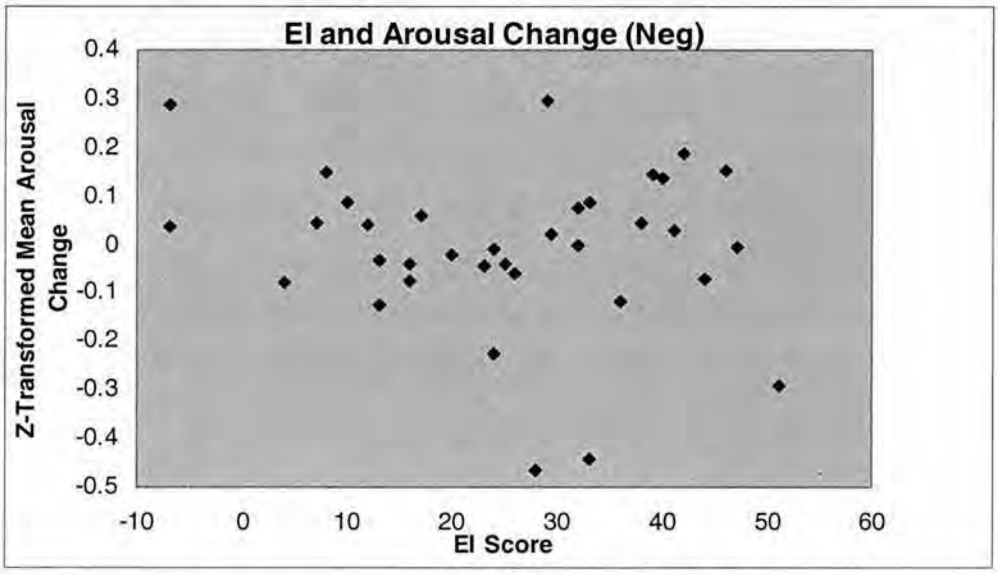
It is also possible with a SCL design such as this to observe two participants, one of whom regulates very well and returns his affect to baseline the other of whom regulates poorly and shows an increase in arousal, to both have equal mean SCL changes. This issue is returned to in the discussion but for now it serves to highlight that this is an unavoidable property of SCL designs.

Initial investigation of the scatterplots for both conditions (Figures 5.2 and 5.3) suggests no correlation between EI and arousal change. This would seem to contradict the current-theory prediction that EI scores will correlate positively with reactivity to the affect manipulation. A closer examination of these scatterplots suggests a correlation between EI score and change in arousal from baseline for those that score just above the mean (cut-off score = 28, $M = 25$) and the disappearance of any correlation after this point. The direction of the correlation between EI and arousal change seems to depend on the affective content of the stimulus. These findings seem to support the hypotheses made based on the *awareness* and *flexibility* conjecture.

Inferential statistics confirm the descriptive and graphical observations made regarding these data. As noted above, a smaller change in arousal from baseline is indicative of a greater extent of affect regulation. There was no overall correlation between EI score and arousal change from baseline in either the positive or negative affect conditions ($\rho = .01$ and $\rho = .00$, all $ps > .5$). As such, the predicted negative association between EI score and arousal change (primacy of 'sensitivity' over 'control') was not supported.

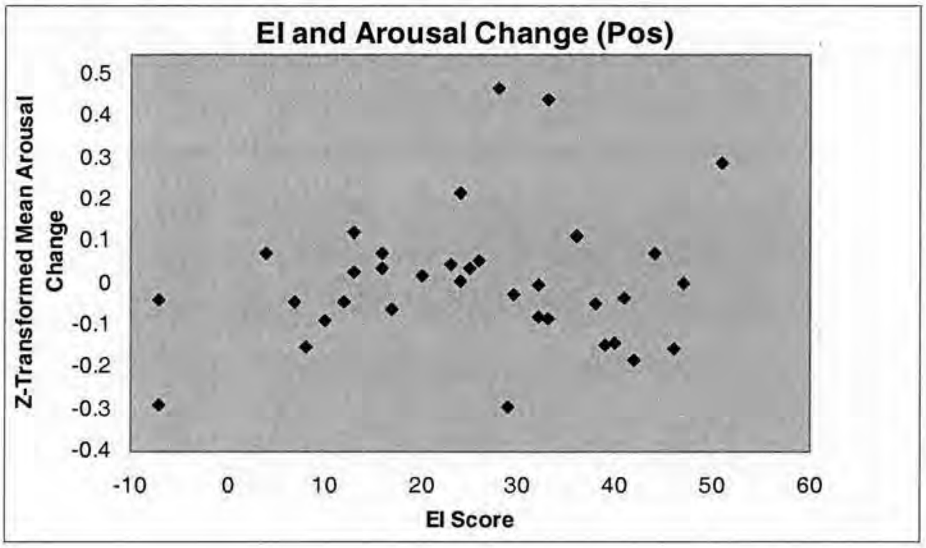
In both affective conditions, EI score was significantly correlated with reactivity to induction and thus affect regulation, but only for those participants who score less than or equal to 28 on the EIS. In the negative condition, EI was negatively correlated with arousal change ($\rho = -.56$, $p < .01$, $N = 22$) amongst low-EI individuals, suggesting that relatively higher-EI individuals were better at maintaining a baseline level of arousal. A similar case was observed in the positive affect condition – EI scores were positive associated with arousal change from baseline ($\rho = .52$, $p < .01$, $N = 22$) suggesting a 'neutralising' effect of EI. The lowest scorers were made happier by the stimuli but participants with EI scores near the mean hardly reacted at all. In both the positive and negative conditions, high-EI individuals appear to regulate their affect more variably, as evidenced by the non-significant correlation between EI and arousal change ($\rho = -.12$ and $\rho = -.13$, all $ps > .3$, $N = 17$) in that sub-group. The low-scoring (i.e., scored below 28) group had less variable arousal responses in the positive ($SD = .11$) and negative ($SD = .10$) conditions than high-EI scorers ($SD = .18$, $.16$, respectively).

Figure 5.2 Scatterplot of Arousal in the Negative Mood Condition as a Function of EI



In other words, those who scored below 28 (M = 25) displayed a reliable 'neutralising' effect of EI – greater EI was associated with smaller changes in arousal. The differences in direction of the association were due to differences in the stimulus content (stimuli which cause negative affect tend to increase arousal, thus 'neutralising' *must* be a decrease in arousal). Those who scored about 28 showed no connection between EI test scores and reactivity to the affect induction. This what the hypothesis predicted: no correlation between EI test score and reactivity amongst high-EI individuals owing to their *flexibility* and *awareness* but this relationship should exist amongst low-EI individuals because they regulate in a reliable manner in line with their EI level.

Figure 5.3 Scatterplot of Arousal Change in the Positive Mood Condition as a Function of EI



5.4 Discussion

This study tested two competing hypotheses. According to current EI theory it was predicted that EI test scores would be negatively related with reactivity to affect induction. Specifically, it was predicted that high-EI individuals would be better at affect regulation and thus that they would react less to any affect induction. This hypothesis was not supported as there were no significant correlations between EI test scores and affective reactivity in either condition ($\rho = .01$ and $\rho = .00$, all *p*s > .5). Regardless of whether the logic of either previous affect regulation study (i.e., Petrides & Furnham, 2003; Ciarrochi et al., 2000) is considered sound it is certain that *some* sort of overall correlation between EI test scores and affect reactivity would be expected based on current EI theory. However, it was not possible to reject the null hypothesis which accompanied the hypothesis based on current EI theory.

The *awareness* and *flexibility* conjecture drives a competing hypothesis. It was predicted that for high-EI individuals, there would be no correlation between EI score and affect reactivity (and thus, affect regulation), because high-EI individuals perceive more subtlety in affective stimuli (e.g., the nuclear explosion could engender feelings of anger, shock, horror, sadness, etc) and behave in a greater variety of ways in response to it. This hypothesis was supported in both experimental conditions ($\rho = -.12$ and $\rho = -.13$, all *p*s > .3, $N = 17$). Low-EI individuals, on the other hand, were expected to manifest essentially what current EI theory would predict: an overall correlation between EI score and reactivity to affect induction. This prediction was also supported in both the positive ($\rho = .52$, $p <.01$, $N = 22$) and negative ($\rho = -.56$, $p < .01$, $N = 22$) conditions. For the sake of clarity, the roughly inverse relationship between reactivity and regulation should be repeated: SCL measures reactivity and less reactivity is interpreted as evidence of greater regulation.

The relevance of these findings to this thesis is immediately apparent, as they seem to support the conjecture. However, this study is relevant beyond the scope of this conjecture as it fails to replicate any kind of overall correlation between EI scores and reactivity and thus affect regulation. This is a finding that runs counter to *ability* theory and because it is the first such EDA-related finding, it is of some importance even considered separately from the *awareness* and *flexibility* conjecture. It is strange that the overall EI-reactivity (i.e., regulation) correlation here was similar to the null finding of Ciarrochi et al. (2000) rather than positive as in Petrides and

Furnham (2003), given the methodological similarities between this study and Petrides and Funrham's (2003). It seems that EI test scores do not correspond with these lab-bound indicators of affect regulation or heightened sensitivity to affect.

As was noted above, SCL is a method with some disadvantages. Most critically, it is not usually possible to ensure that every moment of the stimulus contains the desired affective content. This loss of precision can be corrected by using an SCR paradigm and chapter 6 will investigate what effect this methodological change would make. SCL procedures in combination with z-transformation may have produced anomalous results. Astute readers will have noticed the high degree of similarity between the two scatterplots above. They were almost 'mirror images' of one another, with a strong negative correlation between values in each condition. Because z-transformation uses an 'overall' mean which is derived from the positive, negative, and baseline conditions, it is by definition nearly the exact mid-point between all of a participant's readings and as such the overall 'average' changes in arousal for each experimental condition could be of similar magnitude, if not sign. Between these factors and chance similarity between the stimuli, the similarity of the results of the two conditions begins to make sense. However, even though this overall mean splits the data roughly in half and thus explains some of the similarity between conditions, it doesn't compromise the correlation coefficients. It is hard to conceive of a way in which a simple subtraction and division procedure within a single variable would explain inter-variable covariance.

Also, as will be returned to in the discussion section, the subgroups were created based on an arbitrary point near the mean and although the effects were notable with this cut-off point, the results would have been much different if a different cut-off point had been used. It is probably fair to point out that the purpose of this experiment and indeed this thesis in general was to determine if there was *any* utility in the conjecture. By using the most generous cut-off point it was possible to see if the conjecture functioned under ideal circumstances. Obviously, if the results disconfirm the conjecture-driven hypotheses even under ideal circumstances, the conjecture can be easily discarded. In some ways, the failure of a theory under generous conditions is more informative than its failure under stringent conditions.

A subtler and more interesting point is that this study only investigated un-prompted affect regulation and its relationship to EI, much like the study by Petrides and Funrham (2003). It was previously suggested that these researchers might have arrived at different findings from Ciarrochi et al. (2000) because they used an incidental measure of affective state rather than requiring that participants complete some sort of cognitive task. It is possible that the different behaviours that high-EI people engage in are due to different task demands. More importantly, the language used in the conjecture being tested in this thesis places a great deal of weight on 'flexibility', especially as it pertains to ambiguous stimuli, so it seems critical that this 'flexibility' be tested in some manner. This study showed that when presented with an ambiguous affective stimulus, high-EI individuals behaved in a variable manner and low-EI individuals simply regulated affect as best they can. To complement this information, an experiment which tests the converse must be created to determine how high-EI individuals behave when the stimulus is clear, rather than ambiguous. This is precisely what chapter 6 will test.

Sample generalisability was also an issue in this study as participants were primarily university undergraduates. There were no significant gender differences in SCL in either affect condition, so the probability of gender influencing the results may not be as great as it might seem. Of course, most of the participants were undergraduates, and all of the normal problems of generalising from a student sample to the general population apply here as well. Future research in the general population would be helpful in deciding whether the effects observed here transfer to different samples.

Sample generalisability also impacts these results because of its importance to the sub-group creation technique used here. Median and mean splits are a commonly used technique but these values are sample bound. That is, there is no external check on whether the mean value does in fact estimate the population mean. Because the accuracy of the sub-group splits is so important for the clarity of the findings of these studies, it would be ideal to either split participants according to a normed mean or to use participants from the general population. This issue is partially mitigated because this research is preliminary, but any future research would need to use group splits which are externally justified. An attempt to do so is made in chapter 9.

5.5 Summary

In this study predictions were made based on existing EI theory and the conjecture. It was predicted according to existing EI theory that EI test scores would be positively associated with affective reactivity due to high-EI individuals' greater sensitivity to affect. This hypothesis was not supported. According to the conjecture it was predicted that low-EI individuals would display a 'neutralisation effect' of EI. That is, it was predicted that low-EI individuals would show a correlation between EI test scores and affect reactivity such that relatively higher EI would be associated with decreased reactivity. It was also predicted that high-EI individuals (i.e., above the mean) would show no correlation between their EI test scores and affective reactivity. Both of these predictions were supported. These results seem to contraindicate current EI theory in favour of the conjecture.

Chapter 6: Prompted Affect Regulation and the *Awareness* and *Flexibility* Conjecture

6.1 Introduction

In chapter 5 it was shown that when presented with affective stimuli without any clear task, high-EI individuals will regulate their affect in a more variable manner than low-EI individuals. There was no correlation between EI test scores and reactivity to the affect induction and this seems to suggest that there is no evidence of great affect regulation in high-EI individuals. This appears to be the case regardless of whether they view positive or negative stimuli. This finding is contrary to what was expected based on current EI theory. The predictions made based on the conjecture were more successful, however. However, it was noted that the SCL paradigm used in chapter 5 had methodological weaknesses that could be corrected by using an SCR paradigm. More importantly, the role of task clarity in the variability of high-EI participants' responses was questioned. The purpose of the present study is to use SCR, a more precise technique, to determine whether the findings of the previous study were simply due to the lack of a clear task.

6.1.1 Task Clarity and *Flexibility*

The chief theoretical flaw of the previous study was that it confounded task clarity/ambiguity with the conjecture being tested. Because Petrides and Furnham (2003) might have found completely different results from Ciarrochi, Chan, & Caputi (2000) simply because one study included a task that clearly required affect regulation and one did not, it is necessary that future studies of the take task clarity into account.

In addition to the evidence from current EI theory that indicates this confound is important, it is also relevant to the *awareness* and *flexibility* conjecture. This conjecture revolves entirely around *choice*: chapter 4 showed that high-EI individuals react in a greater variety of ways to ambiguous stimuli than low-EI individuals and Chapter 5 showed that high-EI individuals choose a variety of approaches to situations whereas low-EI individuals employ the same approach across situations, albeit with variations in intensity. In other words, high-EI individuals choose a greater variety of approaches to situations but this choice is only possible in a task which is somewhat ambiguous as to what approaches are possible. If a task clearly only has one solution, it is unlikely that high-EI individuals will select alternative solutions. It is more likely that the importance of *awareness* and *flexibility* disappears in situations with clear task demands and in these situations EI is identical to what current theory would predict. Thus a sort of interaction between task clarity and EI level is predicted: in the clear task condition, high-EI individuals would not be expected to behave more variably and thus there should be no evidence for the conjecture. Rather, in a clear situation it is expected that there will be an overall negative correlation between EI test scores and reactivity to the affect induction (and thus, affect regulation) owing to the greater abilities of high-EI individuals to regulate affect, especially when explicitly requested to do so. The ambiguous condition is included to test to conjecture as the presence of some amount of choice is critical for the conjecture. Similar findings to the previous chapter are expected in the ambiguous condition: a neutralisation effect (i.e., negative relationship between EI test score and reactivity to affect – affect regulation) of EI amongst low-EI individuals which disappears amongst high-EI individuals.

6.1.2 Precision and Paradigms

Another flaw that was discussed in chapter 5 was the lack of precision in SCL paradigms. Simply put, it is not possible to ensure that every moment of a 1-minute long film clip contains the affect that is desired by a researcher. Admittedly, a film clip probably only requires some minimal number of affect-laden 'moments' to induce affect (e.g., it only takes one 5-second grisly scene to be 'disturbing'), but it is entirely unclear what this minimum is. SCR paradigms are much more precise. Instead of a dynamic stimuli static images are presented and it is usually possible to determine which affect is induced by various pictures and how intensely is it induced before experimentation. Images are nearly as effective as films for inducing affect

(Martin, 1990; Westermann, Spies, Stahl, and Hesse, 1996) and image databases such as the IAPS (Lang, Bradley, & Cuthbert, 2005) provide images that have been rated at nearly all points along the positive/negative and intense/mild continua.

In addition to the stimulus precision, SCR also collects data in more precise way. Data are collected in 5-7 second epochs rather than minute-long epochs so it is easier to determine how a participant reacted to a given stimulus. Although real data are noisy compared to the elusive 'standard' SCR waveforms (e.g., Boucesein, 1993), it is still much clearer and easier to interpret a group's averaged SCRs to a set of pictures than to a minute of film. Moreover, the amplitude of SCR is a more rigorous and precise indicator of arousal change than a tonic measure of SCL changes from baseline. Habituation effects are less likely in SCR paradigms if affective stimuli are randomised such that picture order effects are prevented. In sum, SCR is a more precise and rigorous measure of arousal and all subsequent studies will use this technique instead of SCL.

The nature of SCR is also more straight-forward than SCL. A smaller SCR amplitude (i.e., maximum in an 'area of interest' in the waveform) indicates that the participant was not as disturbed by the stimulus as a participant with a larger SCR amplitude. Besides z-transformation, no baseline procedures are necessary and no 'arousal change' values are necessary. SCR amplitude values *are* arousal change values. The only calculations necessary are correlations to determine whether EI is related to SCR amplitude in each affective and experimental condition.

6.1.3 Affective Stimuli

Images from the IAPS (Lang et al., 2005) were used to induce affect. The IAPS database contains a record of subjective ratings of the pictures for their affective valence and the intensity of the affect they induce. For this study, positive, negative, and neutral images were chosen based on how people rated them in terms of affective valence and intensity. More detail is provided below. In addition to subjective rating evidence for the experimental efficacy of affect-laden images, there is substantial physiological (i.e., SCR) evidence that images are effective at inducing affect (e.g., Bradley, 2000). The images used in this study were selected according to their subjective ratings of valence and intensity. In the IAPS, ratings of valence and arousal range from 1(low) to 9(high) with neutral images rated somewhere between 4 and 5. For this study, images rated between six and seven on both continua were used as 'negative' stimuli and images rated between two and three on valence and six and seven on the intensity continuum were used for 'positive' stimuli. 'Neutral' stimuli were those which were rated at approximately one in intensity and four to five on affect. Several images were used to induce each type of affect.

6.1.4 Hypotheses

The discussion above explained how the relationship between EI test scores and reactivity to affect induction might be influenced by task clarity. Reactions to a clear request might be different from reactions to an ambiguous task. It is predicted that in the ambiguous task condition, a similar finding to chapter 5 will be observed: a negative correlation between EI test scores and SCR amplitude – a smaller amplitude reflects a lesser reaction to the induction - for low-EI individuals but no such correlation amongst high-EI individuals. In the clear task condition, current EI theory will hold and thus there will be a simple overall negative correlation between EI score and SCR amplitude. This negative relationship if found will be taken to be indicative of greater affect regulation abilities amongst high-EI individuals.This prediction follows Ciarrochi et al's (2000) finding because this study, like theirs, contains a clear affect regulation requirement. A group-difference (i.e., ANOVA) approach could also be used here, but this approach would prevent investigation of the full range of the EI spectrum

6.2 Method

6.2.1 Design

Skin Conductance Response and scores on the TEIQue (discussed below) *trait* EI test were recorded in two within-subjects experimental conditions: a clear task and an ambiguous task.

6.2.2 Participants

Forty Edinburgh University students (fifteen male and twenty five female) took part in this study for £2.50. Their ages ranged from nineteen to thirty five (M = 24.8, SD = 4.2). Participants were recruited using an advertisement on the Edinburgh University student employment website and although this website is viewable by all students at this university, this was still an opportunity sample. Participants were split into high- and low-scoring groups based on their EI scores.

6.2.3 Measures

The TEIque-SF was used in this study. This is a *trait* EI test (Petrides & Furnham, in press) designed to return a global EI score and fifteen subscale scores although the authors caution against liberal use of the sub-scales, presumably because these sub-scales would be two items in length. It contains thirty items. It is responded to using a 7-point self-report Likert scale anchored at 'strongly agree' and 'strongly disagree'. The responses were fairly reliable in this sample (α = .66) of forty participants.

6.2.4 Apparatus

The same computer and skin conductance equipment as in chapter 5 was used in this study with the exception of the computer program that the participants interacted with. The new program used for this study was written by Paul Stevens. Ten images from the IAPS were selected based on the criteria explained above to constitute the 'positive,' 'negative,' and 'neutral' stimuli. The ten pictures were split evenly and randomly between the two experimental conditions.

6.2.5 Procedure

Participants first completed the TEIque and were then attached to the SCR equipment. The electrodes were affixed to the medial phalanx of the index and middle digit of the non-dominant hand to prevent the signal from being affected or degraded by calluses. Participants took part in both experimental conditions in a randomised, double-blind paradigm. At the beginning of each condition participants were presented with an instruction screen on a touch-sensitive monitor which was used to navigate through instruction screens. These two instruction screens were identical save for the fact that the clear condition instruction screen included an instruction to 'attempt to neutralise any emotions you experience as a result of looking at these images.' Verbal instructions to the participants were for them to 'neutralise' their emotions in the clear condition and to 'do whatever they would normally do' in the ambiguous condition. As noted before, each condition comprised fifteen affective images – five of each of the three affective valences – and these images were presented in a random order for five seconds. After each stimulus presentation there was a random rest period of some length between seven and nine seconds to prevent expectation effects. Thus each participant completed the questionnaire, completed one experimental condition, then completed the remaining condition.

6.3 Results

Data in this study were z-transformed in the same manner as in chapter 5 and only the z-transformed data were analysed. The physiological descriptive statistics presented in Table 6.1 are averages of participants' z-transformed SCR peak amplitudes in a region of interest. Participants' SCR amplitudes are expressed as a number of standard deviations from the mean thus the descriptive statistics in Table

6.1 below are group means of individuals' amplitudes, expressed as standard deviations (σ) from the mean. Again, it should be noted that it is the connection between individuals' SCR amplitude and their EI test scores, not group differences, which are of interest.

In order to extract mean SCR amplitudes some additional procedures (compared to SCL) are required. First, responses in all of the picture categories are averaged across participants so that an 'average' waveform is constructed for all picture categories in both experimental conditions. This average waveform is then used to identify regions of interest for the entire sample. Generally, this region of interest is a two-second period that occurs two seconds after stimulus presentation. After the region of interest is identified, SCR amplitude peaks are identified within this region of interest for each participant and these form the raw data used in further analyses.

Table 6.1: EI test score and z-transformed SCR amplitude descriptive statistics for experimental and affective conditions

	TEIque	Cneg(σ)	Cpos(σ)	Cneu(σ)	Aneg(σ)	Apos(σ)	Aneu(σ)
M	130	0.151	0.087	0.019	0.123	0.122	0.018
S.D.	18	0.234	0.247	0.126	0.278	0.274	0.167

Note: TEIque refers to scores on the TEIque Cneg refers to the negative affect pictures in the clear condition, Cpos to the positive affect pictures in the clear condition, Cneu to the neutral pictures in the clear condition, Aneg to the negative pictures in the ambiguous condition, Apos to the positive pictures in the ambiguous condition, and Aneu to the neutral pictures in the ambiguous condition. Physiological variables are expressed in terms of standard deviations from the mean.

As in chapter 5, visual inspection of the data as well as normality tests indicated that these results were non-normal and as such only non-parametric measures are used here. A Mann-Whitney U test revealed a significant gender difference in SCR amplitude in the ambiguous negative condition ($U = 103.5$, $p < .05$) but no significant gender difference in EI scores.

EI sub-groups were formed by a mean (M =130) split which resulted in a low-group of seventeen (nine female, eight male) participants and a high-group of twenty three participants (sixteen female, seven male). In the ambiguous condition, the EI scores of low-EI participants (i.e., those that scored under the mean) correlated moderately negatively with SCR amplitude in the negative affect condition ($\rho = -.49, p < .05$) but there was no such correlation between EI scores and SCR amplitude in the positive image condition. The correlations between EI test scores and SCR amplitude in the different experimental conditions and different EI sub-groups are presented in table 6.1 below. This relationship is depicted in Figures 6.1 and 6.2 and of note is the manner in which there is a strong 'neutralisation' trend which disappears when the entire sample is analysed.

Table 6.1 Correlations between EI test scores and SCR amplitude in different experimental conditions.

	Z-Transformed SCR Amplitude			
EI Subgroup	Amb. P	Amb. N	Clear P.	Clear N.
EI test score (Overall, N = 40)	0.07	0.11	0.18	0.16
EI test score (high, N = 23)	0.14	-0.10	-0.13	-0.17
EI test score (low, N = 17)	0.04	**-0.49**	-0.02	0.36

Note: Amb. P refers to the ambiguous positive condition, Amb. N to the ambiguous negative condition, Clear P. to the clear positive condition, and Clear N to the clear negative condition. Different rows display correlations in either the overall sample, the high-scoring group, or the low-scoring group. Bold numbers are significant at the $p < .05$ level.

Figure 6.1 The relationship between EI scores and SCR amplitude in response to negative stimuli in the ambiguous condition – low-EI participants only.

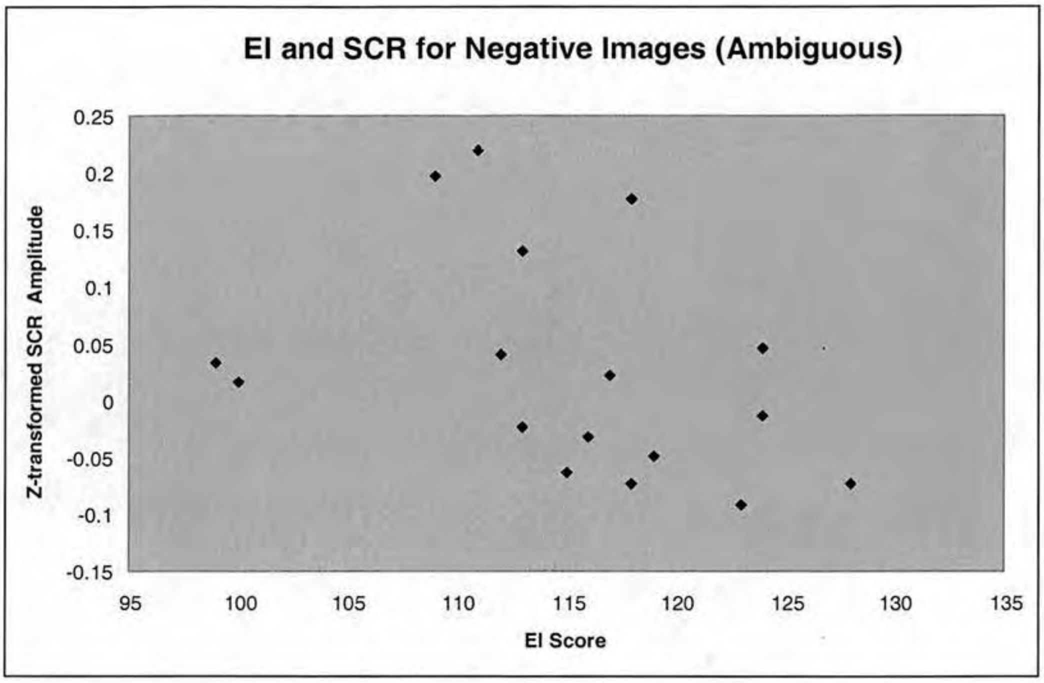
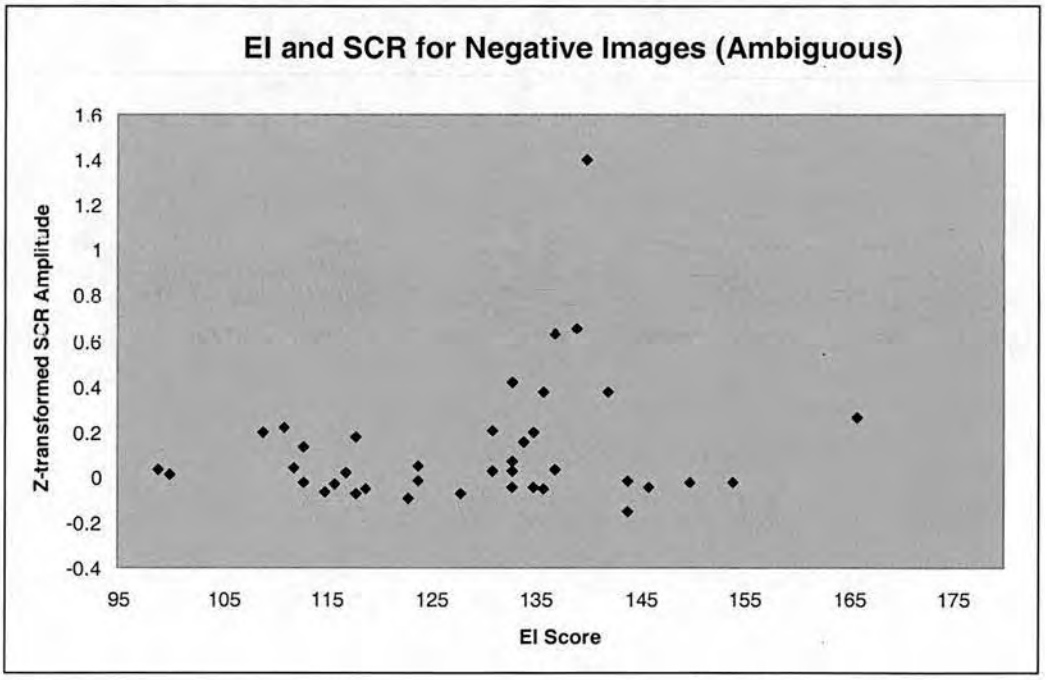


Figure 6.2 The relationship between EI scores and SCR amplitude in response to negative stimuli in the ambiguous condition – entire sample.



Figures 6.1 and 6.2 illustrate the *awareness* and *flexibility* conjecture. Higher SCR amplitude indicates greater anxiety in response to the stimulus so it appears that increased EI is associated with less anxiety (i.e., affect neutralisation) up to a certain point, after which there is no discernable relationship between EI and affective reactions. In other words, given an ambiguous situation, low-EI individuals neutralise their affective reactions in a predictable manner in line with their EI level and high-EI individuals show no predictable behaviour in an ambiguous situation.

These findings partially support the conjecture-driven hypothesis regarding the relationship of EI and affect regulation in ambiguous situations. There were no significant correlations between EI and SCR amplitude in any of the experimental conditions when the entire sample was examined. This finding disconfirms the expectation (driven by current EI theory) that EI scores would correlate positively with SCR amplitude in the clear condition and indicates neither greater sensitivity nor greater regulation amongst high-EI individuals.

6.4 Discussion

A moderate negative correlation between SCR amplitude and EI score was observed for low-EI (i.e., below the mean) participants in the ambiguous negative condition ($\alpha = -.49$, $p < .05$) but no similar correlation in the ambiguous positive condition. There were no correlations between EI and SCR amplitude in any conditions for high-EI. This provides only partial support for the *awareness* and *flexibility* conjecture as it shows that the hypothesis based on current EI theory – that there would an overall correlation between EI and SCR amplitude due to increased 'sensitivity' in high-EI individuals – was not supported as there were no overall correlations between EI and affect regulation in any of the affect or experimental conditions.

There are theoretical and methodological issues in this study. The two most important points cut to the core of this thesis: 1) it is not yet certain that participants are indeed regulating affect when instructed to do so; and 2) the suitability of physiological indicators as a dependent variable must be defended. Once a response has been made to these points, there are a number of interesting theoretical points which could be made about the results of this study with reference to current EI theory and the conjecture being tested in this thesis.

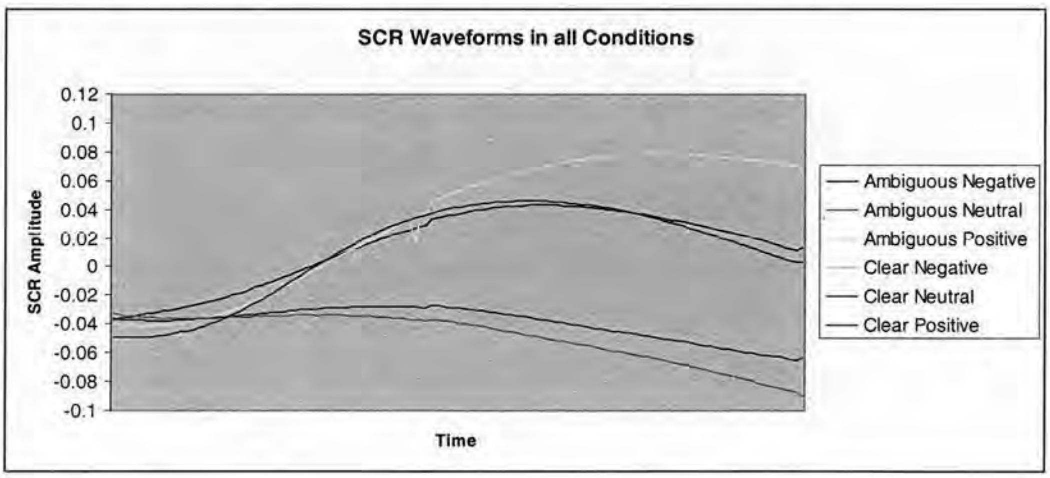
It would be tempting to respond to 1) by simply suggesting that is certain that participants are regulating affect to different degrees because their SCR amplitudes differ from one another (i.e., that we are certain they are doing the task based on their score on it), but this response is a tautology. This response is circular because SCR amplitude is already being used as an indicator of reactivity to the affect induction and thus, affect regulation. In essence, to use this response would be to state 'SCR is being used as an indicator of affect regulation and we know participants are regulating their affect by examining their SCR amplitude,' which would be analogous to a memory researcher suggesting that 'longer rehearsal time is used to define depth of processing and we know participants process more deeply because they rehearse longer.' This difficulty is partially mitigated by the fact that this investigation regards the connection between EI and lab-based affect regulation indices (rather than a theory of what causes better affect regulation) and as such the index of affect regulation may be more arbitrary than, say, a memory task in a theory of memory. However, it is still critical to know that participants are actually regulating affect, otherwise any correlations found here are meaningless. It could also be argued that the *trait* EI measure used here provides an independent measure of affect regulation ability, but this is hardly the same as an index of affect regulation on a specific task. These attempts to avoid the question fail. Rather what is needed is a design which includes two independent measures of affect regulation in the same study. The previous study which used SCL only goes half way towards this aim. In order to bypass the tautology and to be certain that participants are indeed regulating affect, it will be necessary to include two measures of affective state. To this end and for a number of other reasons explained below, the future studies in this thesis will include self-report measures of affective state in addition to SCR.

Another formulation of 1) would suggest that the differences between experimental conditions are due to some sort of confound, but the methodology used here makes this unlikely. The experimental conditions were counterbalanced and differed in only two respects: the images used in each affective category and whether the task was clear or ambiguous. While it is possible that the different images used in the different conditions are responsible for the different SCR amplitudes, this is unlikely as all the images were selected along similar criteria and as there is no alternative to using difference images as the repeated use of identical images would cause habituation. The only remaining difference between experimental conditions was what participants were instructed to do, i.e., the experimental manipulation of task clarity. As was noted above, an independent measure of affect regulation would be necessary to be certain that participants were actually regulating more effectively than in the ambiguous condition and future studies will include such a measure. Still, it is safe to say that the experimental conditions were similar enough to one another that there is little chance that differences in SCR amplitude were due to a confound.

The suitability of physiological indicators can be defended by appealing to the robustness of the method, but subjective indicators of affect might add a great deal to future studies. Some participants complained that it was very difficult to 'neutralise' affect within a five-second period. In order to determine if SCR was sensitive to affect regulation, the waveforms (averaged across the five pictures which formed each affective category) of the participants with the highest and lowest SCR amplitudes (i.e., most or least evidence of affect regulation) in both experimental conditions were examined. These *post-hoc* analyses suggest that there are meaningful differences in how participants respond when they are explicitly told to neutralise their affect. All comparisons of waveforms within the same affect but in different experimental conditions (i.e., positive ambiguous and positive clear) revealed differences in peaks and forms, with the exception of one comparison. Figure 6.3 illustrates this finding. Participants who were 'better' at affect regulation seemed to exhibit greater differences between the clear and ambiguous experimental conditions. In short, the physiological data suggest that participants behave differently in the clear and ambiguous conditions – a successful manipulation. Moreover, the group (and extreme sub-group) data suggest that participants experienced smaller changes in arousal when explicitly told to regulate their affect when they viewed negative but not positive pictures – in other words their physiological responses indicate that they regulated more effectively when they were explicitly told to do so. Thus it appears that the five-second stimulus duration and the SCR paradigm used here were sufficient to accurately measure affect regulation. That said, the affective response is also a subjective phenomenon and as such the inclusion of self-report indices of affect could only serve to improve future studies.

It is also possible that the physiological indicators of affect used in this chapter and the previous chapter are overly simplistic. Even if the breadth of affective response is ignored and only physiological aspects or indicators of affect are examined, there is still considerably more to the physiology of affect than skin conductance. Activity in the CNS and other peripheral organs also accompanies affective response and although SCR is a good measure of arousal, it is relatively insensitive to affective valence or complex emotions. Thus it is possible to criticise this study for over-simplifying the physiological aspects of affective response. This issue is examined in more detail in the general discussion.

Figure 6.3 Average SCR waveforms for each condition and affective category. Measurement of SCR amplitude is in microsiemens.



In order to circumvent the tautology discussed previously and also to cohere better with previous studies, further studies will include self-report indices of affect in addition to physiological indicators. In similar prior studies (i.e., Ciarrochi et al., 2000; Petrides & Furnham, 2003) subjective measures of affective state were used and in this thesis, only physiological, more 'objective' measures have been used. This methodological difference may explain why previous studies found overall correlations between EI scores and affect regulation and the current study did not. It may be that previous findings were due to some kind of self-report bias which is not present in this study. The purpose this thesis was to investigate more 'objective' indices of affect and how they relate to EI test scores, but this aim is not undermined by the use of self-report measures. A replication which uses self-report and physiological methodology would be helpful to understand the reasons for the null findings here. It would also improve the ecological validity of this study. Of course, self-report measures could not be used both as a manipulation check and a hypothesis test as this would require that they perform 'double duty' and would be as tautological as using SCR for both duties.

It is unclear why the prediction based on the conjecture was only supported in one affective condition. Based on the conjecture, there is no reason to suspect that the connection between EI and affect regulation would be different when viewing different types of stimuli. Of course, there is a vast difference between 'happy' and 'anxious,' from biology to behaviour, so it is possible that there is some complex interplay between affective stimuli and the way *awareness* and *flexibility* function in the conjecture. Perhaps high-EI individuals are simply not predictable in how they react to positive stimuli regardless of whether they are given a clear or an ambiguous task. Perhaps the findings in chapter 4 were a fluke. Perhaps there is something unique about positive stimuli that might alter the relationship between EI and SCR amplitude. The present study is insufficient for any conclusion on the matter and further replications will be necessary.

Another troublesome issue is the lack of any overall correlation between EI score and affect regulation. Although similar null findings were observed in some of the comparisons carried out by Ciarrochi et al., (2000), current EI theory would certainly predict *some* kind of connection between EI test scores and the degree of affect regulation, whether it be negative (indicating that high-EI is associated with greater 'sensitivity') or positive (indicating that high-EI is greater 'regulation'). The results from this study seem to disconfirm current EI theory. Simply put, there does not seem to be a correlation between EI test scores and individuals' physiological reactions to affective stimuli and it is very difficult to reconcile these findings with present EI theory. There is ample reason to expect that (when presented with a clear task) high-EI individuals would perform better at it than low-EI individuals but this study and the prior study show that this is not the case. It is possible that EI does not function on the physiological level and the following study will replicate the current study to determine if this is the case. It will also include a preliminary, incidental probe into the connection between EI test scores and subjective indices of reactivity to affect induction. However, because physiological reactions are so important to many emotions, it seems inappropriate to suggest that physiology is out of the remit of EI. Future studies will be necessary to determine if these findings are robust but it does seem that, however tentatively, current EI theory does not make accurate predictions about physiological indices of affect regulation. The *awareness* and *flexibility* conjecture receives very little support as well.

Of course, these results only reflect a failure of EI tests to predict a very specific physiological indicator of the affective response. It is possible that EI tests predict other aspects of the physiological response to affective stimuli (e.g., heart rate, pupil dilation, etc.) better. As was noted above, EI tests may also predict subjective measures of affective response reliably. The issue of how convincing or generalisable these findings are is taken up in more detail in the general discussion.

It may also be the case that EI operates at different 'levels.' That is, perhaps some EI-related skills function at a conscious level whilst some function at a more fundamental or biological level. Just as performance on inspection time tasks is more strongly correlated with fluid ability measures (Deary, 2000; Austin, 2004) than crystallised ability measures, perhaps some branches of EI are more heavily correlated with some lab-based tasks than others. There have been some studies (e.g., Austin, 2004; Farreley & Austin, 2005) which have substantiated this claim but it is not presently part of the core of existing EI theory. Additionally, the power of this defence relies on the clarity of EI factor structure which was shown in chapter 4 to be somewhat suspect. Still, if the psychometric hurdles can be bypassed, it may be that differences in predictive validity between EI factors may account for the findings of this study.

As in the previous studies, gender effects are worth discussing. For some reason (many possibilities present themselves) there were more female than male participants. As was noted above, there was a gender difference in reactivity to the affect manipulation but there was no significant difference in EI scores between genders and the gender composition of both EI sub-groups was approximately equal. It might seem tempting to attribute the findings of this study to a gender effect, but it must be recalled that the hypothesis, and even the conjecture itself, has to do with how EI skills are used, not whether one sub-group is 'better' than another. When men and women were analysed separately *post hoc,* neither gender showed any evidence of an overall correlation between EI test scores and SCR amplitude regardless of gender (all p's > .2) so it is unlikely that gender differences are responsible for the findings here.

It might be argued that inducing affect in the lab lacks the ecological validity of other methods. This may be the case, but as was noted there is evidence to suggest that using images or films in the lab is an effective means of inducing affect that avoids demand characteristics (Martin, 1990; Westermann et al., 1996). So although it is probably true that there are more ecologically valid ways of inducing affect, the methods used here are probably acceptable on a number of other grounds: precision, replicability, objectivity. Unfortunately, it is usually necessary to sacrifice ecological validity for these other desired qualities, and as was explained in chapters 1 and 2 this thesis is focused specifically on physiological indices of affect as they seem to be the most precise, replicable, and objective (ignoring the multiple ambiguities in this term) of affective measures. In short, it is true that this study lacked some ecological validity, but it also benefits from some qualities that a more ecologically valid study would lack.

There are numerous replications and extensions which could be performed on this study. Primarily, any replication could use self-report as well as physiological indicators of affect in order to address the ecological validity critique and to determine if the lack of an overall correlation was due to the methods used here. It might be useful to employ non-IAPS images as a means of inducing affect – the Ekman faces would probably suit this purpose equally well and might be argued to be more appropriate as they all pertain to facial displays of emotion. It also might be illuminating to see how people with different EI levels responded to some of the more violent or disturbing images in the IAPS, but such a study may be considered unethical or at least unpleasant by some. There are also numerous ways in which a series of experiments would take a form only vaguely similar to the series of experiments in this thesis: different physiological indices, a different EI-related skill, etc. More space will be devoted to discussion of future directions in the general discussion section below.

It is also possible that the EI test used in this study caused the null findings. In chapter 5, the 41-item EIS (Austin, Saklofske, Huang, & McKenney, 2004) was used whereas the TEIQue short-form (Petrides & Furnham, in press) was used in the present study. The results of chapter 5 also supported the conjecture more clearly. Although these two tests are similar in that they are both *trait* EI tests, they differ in their factor structures, length, and construction. Given the previous evidence that different EI tests can result in notably different results in two similar studies (e.g., Austin 2004; 2005), it is possible that something similar is occurring here. However, because this study was the first SCR study of EI and affect regulation it would seem prudent to determine if these findings are reliable before replicating with a different EI test.

6.5 Summary

In this study, participants completed the TEIque, a *trait* EI test, and viewed a series of affect-laden pictures while their SCR was recorded in both an experimental procedure with a clear or ambiguous task. Based on the *awareness* and *flexibility* conjecture, it was hypothesized that when presented with an ambiguous task, the findings of chapter 4 would replicate – that the EI scores of low-EI individuals would correlate with their affect regulation but that this correlation would be not be present for high-EI individuals. It was also predicted, based on current EI theory, that when given a clear task, there would be an overall negative correlation between EI and reactivity to the affect induction and thus a positive relationship between EI test score and lab indicator of affect regulation. The first hypothesis was partially supported but the second was not. The importance of these findings, especially the null overall correlation, was discussed with reference to present EI theory. It was noted that an important replication would include both self-report and physiological indices of affect and it is to this replication that we now turn.

Chapter 7: Prompted Affect Regulation and the *Awareness* and *Flexibility* Conjecture – Self-report and Physiological Indices

7.1 Introduction

The results of chapter 6 provided partial support for the *awareness* and *flexibility* conjecture, but this partial support was tempered by the lack of subjective reports of reactivity to the affect induction. Prior studies of EI and affect regulation (e.g., Ciarrochi, Chan, & Caputi, 2000; Petrides & Furnham, 2003) have used self-report methods to measure affect and thus affect regulation. The studies in this thesis have only used physiological measures. These lab-based, physiological measures of affect should be complemented by subjective measures in order to increase ecological validity and to serve as a manipulation check.

A subjective measure of affect could take a variety of forms, but the most relevant studies (i.e., studies of EI and affect regulation) use Likert scales completed by either the participant (Petrides & Furnham, 2003) or a group of observers (Ciarrochi et al., 2000). Other, non-EI related studies of affect (e.g., Deiner, Larsen, Levine, & Emmons, 1985) often use similar Likert scales. Although these non-EI studies of affect often use these self-report measures as 'manipulation checks' rather than solely as dependent variables, they are still clearly used to measure affect. This approach has intuitive appeal: the best way to know to what a person is thinking is probably to ask him/her.

The inclusion of the subjective measure of affective state serves a number of purposes. Most importantly, this measure will remove the tautology discussed in the previous chapter. The self-report measure can be used to ensure that participants are indeed regulating their affect whilst the physiological measure is used to determine how effective they are at regulation (i.e., the degree to which they react to the stimuli at all). The self-report measure will also add ecological validity to the study in chapter 6 and will increase sensitivity to attempts at affect regulation because this measure will be completed at the participant's own speed and thus will not be constrained by the five-second stimulus presentation time. This methodological improvement will also make it possible to determine if the correlation between EI and performance on lab-based emotion-related tasks depends on the outcome measure used. Simply put, including a self-report measure will allow a great deal of insight that would not otherwise be possible.

In this study, the subjective indicators were not used to test the same hypotheses as the physiological data because the subjective reports were not 'change' scores. Also, as is noted below, the self-report data were included as a manipulation check and it is not appropriate to use the same measure as a manipulation check and as a dependent variable. It would not be appropriate for the self-report data to be used both as a manipulation check and a hypothesis test. To do so with the self-report data would encounter the same tautology that their inclusion was designed to prevent! The self-report hypotheses should thus be treated with care and perhaps are best considered an incidental, superficial probe of whether or not subjective data mirror the physiological data.

The predictions for the self-report data differ slightly from the predictions for the physiological data. SCR amplitude is a measure of a participant's change in arousal from baseline and thus is an index of his/her reactivity to the affect induction. The self-report indices were static ratings of the images in terms of their intensity and valence and thus do not represent any change in arousal. However, these static indices of affective state can still be used to make inferences about affect regulation. That is, although they do not control for baseline affective state, it is still informative to know if EI test scores are associated with greater or less negativity (or positivity) in ratings of affect in response to negative or positive stimuli. It would be expected based on current EI theory that because of their greater affect regulation (or repair) abilities, high-EI individuals would rate their affective state as more positive after viewing negative stimuli. Ciarrochi et al. (2000) made similar predictions. Specifically, current EI theory would predict a positive association between EI test score and valence ratings in response to the positive and negative pictures clear condition – reflecting the positive impact of affect regulation. Current EI theory would also prompt us to expect that because of their heightened affect regulation abilities higher EI would be negatively associated with ratings of the intensity of the affect induction in the clear condition. As was noted before, due to the importance of choice in the operation of the *awareness* and *flexibility* mechanisms in the conjecture, predictions based on the conjecture are only made for the ambiguous condition. The conjecture leads to the prediction that in the ambiguous condition, amongst low-EI individuals there would be a negative correlation between EI test score and ratings of intensity for positive and negative images. It would also be expected that that amongst low-EI individuals in the ambiguous condition there would be a positive association between EI test score and valence ratings for positive and negative images. In other words, the conjecture would predict that in the ambiguous condition, the results from the low-EI individuals' responses would be similar to what current EI theory would predict for the entire sample in the clear condition. The conjecture would predict no correlation between EI test scores and self-report indices of affect amongst low-EI individuals in the clear condition regardless of stimulus valence.

The investigation into the self-report data should be read as an interesting aside rather than the main thrust of this study. These measures were included to ensure that participants were actually regulating affect when requested to do so and also to break the circular argument that characterised the interpretation of chapter 6. Simply, two independent measures were needed: one to ensure that the manipulation was effective and one to test the hypotheses. It would be inappropriate to use the self-report measures for both purposes. However, it would also be a waste to not examine the potential correlations between EI test scores and reactivity to the affective stimuli, at least incidentally.

The physiological predictions for this study are identical to the previous study. The conjecture leads to the prediction that when the emotional task is left ambiguous, there will be no correlation between EI test scores and reactivity to the affect induction (i.e., SCR amplitude; affect regulation) amongst high-EI individuals whilst low-EI individuals will display a negative correlation between EI score and SCR amplitude (i.e., a correlation between EI scores and affect regulation effectiveness). Current EI theory leads to the prediction that in the clear task condition, EI scores will correlate negatively with SCR amplitude in the entire sample indicating that high-EI individuals will be less reactive to emotional stimuli owing to their heightened affect regulation abilities. Excepting the inclusion of the self-report measure and a smaller number of images for each affect category (three each for neutral, positive, and negative as opposed to five) this study is a replication of chapter 6.

7.2 Method

7.2.1 Design

After completing a *trait* EI questionnaire, SCR was recorded while participants viewed positive, negative, and neutral images in two experimental conditions (clear and ambiguous task) and rated each of these images on two scales: one with 'positive' and 'negative' poles and one with 'mild' and 'intense' poles in a within-subjects, correlational design.

7.2.2 Participants

Forty students were recruited through the student and graduate employment service at Edinburgh University. However, because some participants were not native English speakers, incomplete questionnaire responses, and because there was a technical malfunction which resulted in several participants' data being over-written, only thirty-three were retained. These participants ranged in age from seventeen to thirty eight (M = 22.8, SD = 5.4) and twelve were male while twenty-one were female. Participants were paid £2.50 for their participation.

7.2.3 Measures

The TEIQue was used in this study, as in chapter 6. The reliability of the responses in this data set was high ($\alpha = .87$). The self-report measures which were presented to participants to measure affective state were two 9-point Likert scales, one anchored at 'happy' and 'sad' and one anchored at 'mild' and 'intense'.

7.2.4 Apparatus

The same apparatus were used in chapters 5 and 6 save for the computer program and the picture stimuli. The program was designed by Dr Paul Stevens. In order to keep experimental time at a minimum, three images were used for each affective category (i.e., three positive, negative, and neutral images) in each experimental condition instead of five as in chapter 6. The images used were selected based on similar criteria as in chapter 6. Subjective reports of valence and intensity were keyed in using a touch-sensitive screen identical to the screen used to navigate through the experiment in the previous studies.

7.2.5 Procedure

The procedure in this study was similar to that of the previous study. Participants first completed the same EI questionnaire and were subsequently connected to the SCR electrodes and then were informed about the experimental procedure. As in chapter 6, participants took part in both of two nearly identical experimental conditions in a counterbalanced design. In each experimental condition, participants viewed nine randomly presented pictures and were asked to rate each picture on two scales: positive/negative and mild/intense. After each stimulus, each participant reported how the image made him/her feel on the valence and intensity Likert scales. SCR was recorded during the entire session and SCR data were z-transformed as in chapters 5 and 6.

7.3 Results

Table 7.1 displays the descriptive statistics for the self-report measure of affective state in all of the experimental conditions. As was noted above, the self-report data were primarily used as a manipulation check. The manipulation seems to have worked, as the negative pictures prompted ratings closer to 1 (i.e., extremely unpleasant) than did the neutral or positive images, and the positive images prompted ratings closer to 9 (i.e., extremely pleasant) than did the negative or neutral images. In the ambiguous condition, ratings of valence for the positive and negative pictures differed notably from the ratings of neutral pictures. Ratings of intensity in the ambiguous condition differed less dramatically between the various picture categories than did ratings of valence, but all comparisons of valence and intensity between positive/negative pictures and neutral pictures were significant at $p < .01$, with the exception of the comparison between positive and neutral intensity. This seems to indicate that even in the ambiguous condition, the positive and negative images had the desired effect on ratings of affect. The intensity ratings were expected to be higher (i.e., more intense) in the ambiguous condition and this seems to be the case. Positive and negative images were expected to be rated as more intense than neutral images in either condition but were not expected to differ from one another. These expectations were largely borne out with some exceptions.

Table 7.1 Descriptive statistics for self-report indices of affect

Task Condition		Affective Condition: Positive	Negative	Neutral
Ambiguous	Valence	6.3 *(1.0)*	1.9 *(1.0)*	4.1 *(1.2)*
	Intensity	4.7 *(1.7)*	3.7 *(1.3)*	4.5 *(1.2)*
Clear	Valence	5.6 *(1.0)*	2.5 *(.8)*	4.1 *(.6)*
	Intensity	3.9 *(1.8*	3.1 *(1.8)*	2.6 *(1.4)*

Note: Numbers in parentheses are standard deviations.

Ratings of valence and intensity in the clear condition were less extreme than in the ambiguous condition. There are significant differences (i.e., all p's $< .01$) between ratings of the positive and negative images in the clear and ambiguous conditions indicating that positive ratings were less positive and negative ratings less negative in the clear condition. This suggests that the explicit instruction to 'neutralise' affect in the clear condition was effective.

The subjective data were also included to test the hypotheses driven by the conjecture or current theory using more ecologically valid measures. As was noted previously, the self-report findings should only be taken as an incidental test of the hypotheses. Their primary role was as a manipulation check and it is not reasonable for them to perform 'double duty.' They are included in this and the following chapter for reference. In the clear condition, EI test scores were correlated positively with valence ratings for negative pictures ($\rho = .38$, $p < .05$ one tailed) which indicates that participants who scored highly on the EI test rated the negative images more positively. There were no other significant correlations between EI scores and ratings of valence or intensity in the clear condition. Thus the prediction that higher EI would be associated with a more positive, less intense impact of affective stimuli was partly supported. The correlations between self-report measures of affect and EI test scores in the different experimental conditions are presented in table 7.2 below.

Table 7.2. Correlations between EI test scores and self-report measures of arousal in different experimental conditions.

	Self-Report Rating							
	Intensity				Valence			
EI Subgroup	Am.P	Am.N	C.P	C.N	Am.P	Am.N	C.P.	C.N.
EI test score (Overall, $N = 33$)	-0.10	-0.23	0.16	0.11	-0.06	0.00	0.06	0.38
EI test score (high, $N = 17$)	-0.11	-0.25	**0.52**	0.16	-0.01	-0.14	0.15	**0.53**
EI test score (low, $N = 16$)	**-0.47**	-0.13	0.04	0.06	-0.22	0.10	0.09	0.22

Note: Am. P refers to the ambiguous positive condition, Am. N to the ambiguous negative condition, C. P. to the clear positive condition, and C. N. to the clear negative condition. Different rows display correlations in either the overall sample, the high-scoring group, or the low-scoring group. Bold numbers are significant at the $p < .05$ level.

To assess the *awareness* and *flexibility* conjecture with the self-report data, the sample was split into high-(N = 17) and low- (N = 16) scoring EI subgroups, split at the mean (M = 146). In the low-scoring group, EI scores were correlated negatively with intensity ratings of the positive pictures ($\rho = -.47$, $p < .05$, one tailed). This is the neutralisation effect that was expected with low-EI individuals reporting being more positively affected by the stimuli but middling scorers reporting less impact (i.e., more neutral) of the stimuli. In the high-scoring group there was a positive correlation between EI test scores and ratings of intensity in the clear positive condition and ratings of valence in the clear negative condition. It is strange that the relationship is positive regardless of the type of affect induced, but in any event it is the correlation between EI test scores and reactivity in the entire sample that is of interest, not the high-group only. These results seem to indicate weak support for the conjecture but do not allow rejection of the null hypothesis in favour of current EI theory. Again, these self-report values should be taken as preliminary, incidental tests as their primary purpose was to serve as a manipulation check.

SCR amplitude data were extracted using a similar procedure to what was done in chapter 6. Figure 7.1 displays the average waveforms in each of the picture categories in each experimental condition. This waveform displays a negative trend which is common to SCR responses even after z-transformation. This negative trend does not adversely impact the data as it is individuals' amplitudes that are of interest, not group mean responses. The region of interest remains unchanged: approximately two to four seconds after stimulus presentation.

Figure 7.1 Average waveforms for experimental conditions for the entire sample across the entire range of measurement.

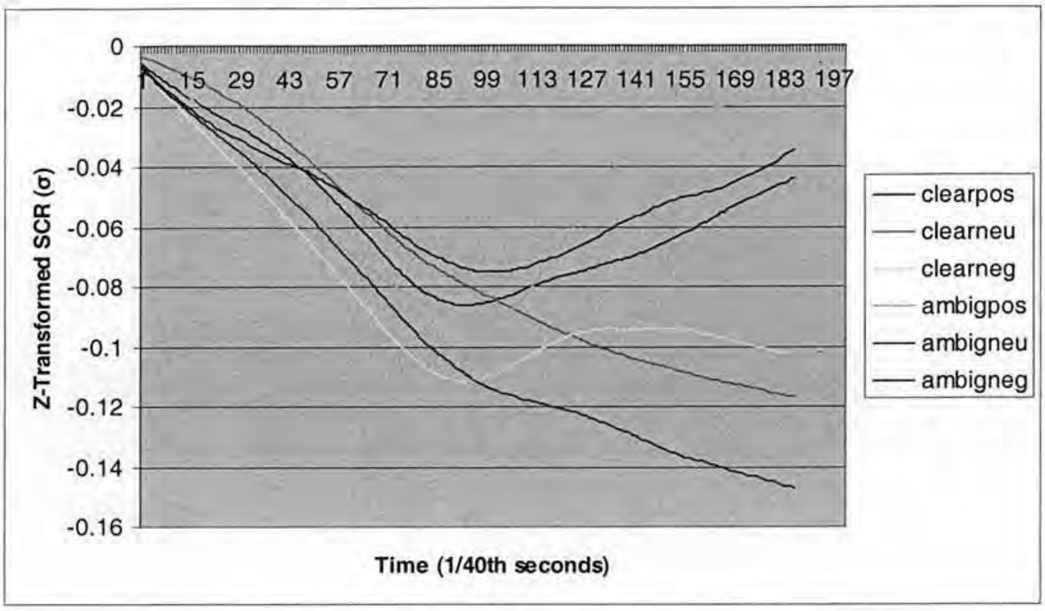


Table 7.3 displays the descriptive statistics for the physiological indices of affective state. Again, these are presented for reference. The low and negative values are due to a relaxation trend in all conditions. Regardless of condition, participants displayed decreased arousal in response to the successive stimuli and this was not due to order as the images were displayed in a random order. The self-report data indicate that the manipulation was effective so it would be redundant (and tautological) to use the physiological data as a manipulation check as well. As in chapter 6, the data in Table 7.2 are averages of individual amplitudes after z-transformation, thus they are expressed in terms of standard deviations from the individual's overall session mean.

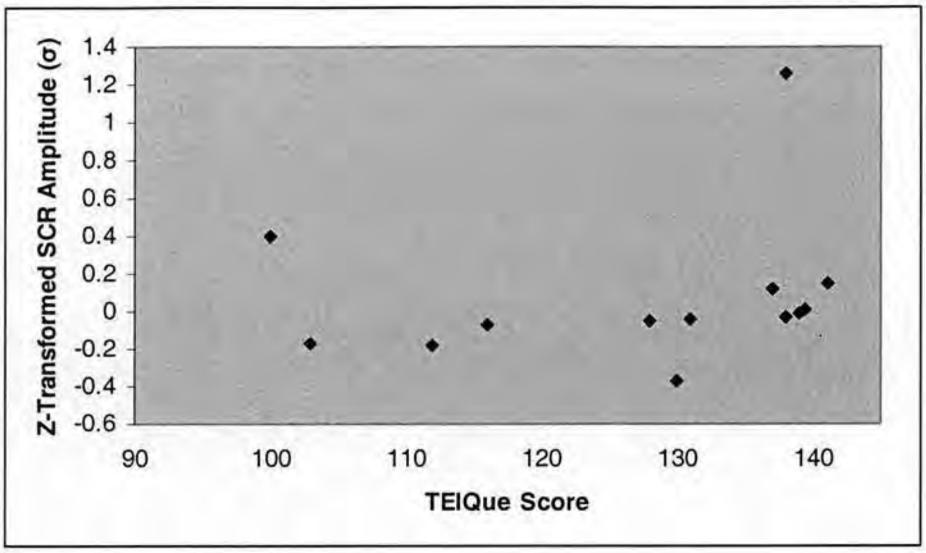
Table 7.3 Descriptive statistics for physiological measures of affect

		Affective Condition		
Task Condition		Positive	Neutral	Negative
Ambiguous	Mean	0	-0.04	-0.07
	S.D.	0.29	0.28	0.25
Clear	Mean	-0.08	0.02	-0.09
	S.D.	0.28	0.27	0.29

Note: All values are expressed in terms of standard deviations off the mean (σ).

The physiological data provided partial support for the hypothesis based on the *flexibility* and *awareness* conjecture. As with the self-report data, the sample was split into high- and low-scoring sub-groups with a cut-off score near the mean (142; $M = 146$, N's = 20, 13, respectively). A moderately strong positive correlation between EI scores and SCR amp in the ambiguous positive condition was observed ($\rho = .47$, $p < .05$) but no significant correlation was found in the ambiguous negative condition ($\rho = .34$, $p > .1$). In the high-EI group, a significant correlation between EI score and SCR amplitude was observed in the clear positive condition ($\rho = -.41$, $p < .05$) indicating that higher EI was associated with weaker reactions to the positive stimuli. This finding supports the hypothesis made according to current EI theory. No significant correlations were observed in the clear condition when the entire sample was analysed. These findings are illustrated in figures 7.2 and 7.3, which show, just as in the previous study, a relationship between EI scores and SCR amplitude amongst low-EI individuals in the ambiguous condition which disappears in the entire sample. A small SCR indicates a smaller increase in anxiety, so figure 7.2 suggests that very low-EI individuals were less reactive to the stimuli whilst mean scorers are not affected at all – i.e., they have 'neutralised' any effect of the stimulus. As in the previous study, this 'neutralisation' effect is not present in the entire sample. The correlations between EI test scores and SCR amplitude in the various conditions are presented in table 7.4 below.

Figure 7.2 The relationship between EI scores and SCR amplitude in reaction to positive stimuli in the ambiguous condition – low-EI participants only



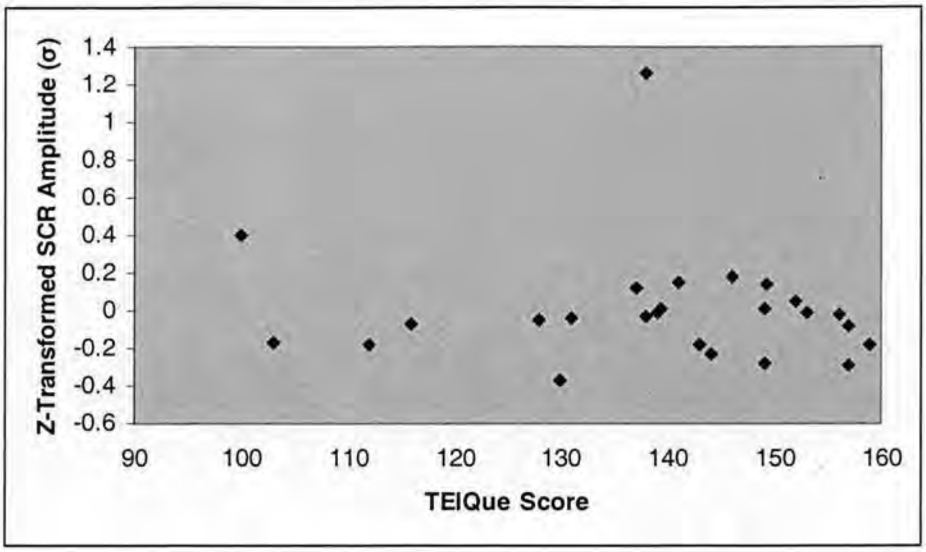
This provides partial support for the conjecture-driven hypothesis as it suggests that low-EI individuals neutralise their affect in accordance with their EI whilst high-EI individuals do not predictably neutralise affect. Partial support was also provided for current EI theory.

Table 7.4. Correlations between EI test scores and SCR amplitude in the different experimental conditions.

	Z-Transformed SCR Amplitude			
EI Subgroup	Amb. P	Amb. N	Clear P.	Clear N.
EI test score (Overall, $N = 33$)	0.01	-0.16	-0.17	0.1
EI test score (high, $N = 20$)	0.09	0.23	**-0.41**	0.14
EI test score (low, $N = 13$)	**0.47**	0.34	-0.21	0.16

Note: Amb. P refers to the ambiguous positive condition, Amb. N to the ambiguous negative condition, Clear P. to the clear positive condition, and Clear N to the clear negative condition. Different rows display correlations in either the overall sample, the high-scoring group, or the low-scoring group. Bold numbers are significant at the $p < .05$ level.

Figure 7.3 The relationship between EI scores and SCR amplitude in reaction to positive stimuli in the ambiguous condition – entire sample.



7.4 Discussion

The self-report results suggest that that all of the manipulations in this study (i.e., affect and task clarity) were effective. In the clear condition EI scores were correlated positively with negative picture valence ratings ($\rho = .38$, $p < .05$, one tailed). This suggests that higher EI is associated with more positive ratings of negative pictures and thus is suggestive of some sort of affect regulation present amongst high-EI individuals that was not amongst low-EI individuals. In the ambiguous condition EI scores were correlated negatively with positive picture intensity ratings ($\rho = -.47$, $p < .05$, one tailed) amongst the low-EI scorers, suggesting that amongst low-EI individuals, increased EI is associated with decreased intensity of positive affect. High-EI individuals displayed no correlations between their EI scores and their self-report ratings of affective state. Given that there were four potential correlations between EI and affective ratings per experimental condition (i.e., positive and negative valence and intensity) to find only correlation is only weak support for either the conjecture or current EI theory. However, the self-report data were mainly to be used as a manipulation check and these hypotheses test results should be taken as a reference only.

The physiological data were the data used to test the hypotheses. These data indicate partial support for the conjecture-driven hypothesis in the form of a significant correlation between low-EI participants' test scores and SCR amplitude ($\rho = .47, p < .05$) in positive ambiguous situations. There were no significant correlations between EI test scores and SCR amplitude when the entire sample was analysed. This seems to contraindicate current EI theory as it suggests that despite clear instructions, there is no connection between EI test scores and this particular indicator of reactivity to affect induction (and thus, affect regulation). However, EI test scores *were* correlated with SCR amplitude in the clear positive condition amongst high-EI individuals only ($\rho = -.41, p < .05$). This unexpected finding suggests that given a clear task, high-EI individuals do regulate affect in line with the EI test scores. It also suggests that removing ambiguity in a task results in high-EI individuals regulating in line with their EI test scores.

Experimenter demand characteristics may seem to have been an issue in this study. The experimental manipulation requested that participants either regulate their affect (the exact wording was 'try to neutralise your emotions') or to simply 'do whatever you would normally do.' This is obviously an unsubtle manipulation and it could be argued that this instruction was an experimental demand characteristic and thus the method was flawed. However, this is unfair. Giving the participants clear instructions is hardly poor experimental design, in fact, it is normally a necessary requirement of a well-designed study. Without clear instructions, participants would have no idea what to do in the study and it would hardly be a useful experiment if the participants were not asked to do anything! It is *not* a flaw (demand characteristic or otherwise) when participants follow experimental instructions. A demand characteristic refers to situations in which experimenters give unintended cues to participants about how to behave or when the manipulation was actually not effective at all but participants unconsciously behave in a way that leads the experimenter to believe it was an effective manipulation. Neither of these situations applied to this study. The participants were given clear instructions and the manipulation check merely shows that they followed these instructions. An obvious manipulation is not problematic either, even in a within-subjects design, unless deception is critical to the design. In this study, participants were clearly told what to do and there was no deception, so if they 'caught on' to what was being asked of them and unconsciously tried harder to perform what was being asked of them, this would merely reinforce the instructions.

At this point two experiments have only yielded partial support for the *awareness* and *flexibility* conjecture, and although one experiment and one psychometric study supported it, it may be appropriate to abandon this theoretical possibility. In this study, the critical condition was the ambiguous task condition and the predictions made about this condition were not well supported by the self-report data. Although the predicted null correlations between EI scores and reactivity (i.e., affect regulation) in the high-EI subgroup were found, it was only possible to reject the null hypothesis in favour of positive predictions in one of four possible comparisons. It would probably be dubious to build a theory on a collection of null findings, as nulls often defy more detailed explanations. Simply put, these null findings alone are insufficient to warrant endorsement of this conjecture without any sort of positive findings to accompany them.

Current EI theory leads to the supposition that if there is any kind of link between EI scores and reactivity to affect induction (and thus affect regulation ability) it should be visible when participants are explicitly told to regulate. This study found that EI scores are only correlated with valence ratings of negative pictures in the clear condition but not with valence ratings of positive pictures or intensity ratings of either. The direction of the correlation found here was identical to that of Ciarrochi et al. (2000) which also involved a clear task. This is probably due in some part to the methods used here, but it seems unreasonable to completely dismiss these findings. Both the physiological and subjective data failed to strongly support the hypotheses based on current EI theory and the conjecture fared only slightly better.

The self-report findings seem to have been of use in this study in that they served as a manipulation check and thus allowed the physiological indices to be used solely for hypothesis testing. Although the self-report findings could be examined as a hypothesis test, any conclusions drawn from these data should be viewed as highly tentative. It would probably be best if these self-report findings were not used as hypotheses tests, however, as they do not in all cases match the physiological findings and as such might confuse the matter. In some ways, the results of this study provide approximately the same amount of support for EI theory as Austin's (2004;2005) inspection time work, so if one is inclined to interpret those findings (reviewed in chapter 2) as support for EI, then these studies ought to be interpreted the same way. It may be more conservative to interpret these studies (and their respective null findings) as evidence in opposition to EI, rather than in support of it. Regardless, a replication of this study would be necessary to draw a firm conclusion about the either the conjecture or current EI theory. Again, the physiological measures should be the data used for any hypothesis tests.

The physiological results prompt the same questions as the self-report results, and are largely similar. The *awareness* and *flexibility* conjecture made two non-null predictions and one was supported. Additionally, present EI theory prompted two non-null hypotheses, neither of which was supported. Thus the physiological results of this study seem to favour the conjecture over than present EI theory, but a replication is in order.

This study improved on the prior study by including a self-report index of affective state in addition to the physiological measure but even with this addition it is still unclear whether the conjecture of current EI makes more accurate predictions. Because there was partial support for both theories (depending on the outcome measure) it seems sensible to replicate this study in order to determine if these results are anomalous. It may also be the case that the EI test used in chapter 5 (44-item EIS; Austin et al., 2004) had a different pattern of associations than the test used in chapters 6 and 7, so it is probably sensible to use the EIS in the replication. If the conjecture is in some way 'bound' to the EIS, it will be clear from the results of this next study. Again, just as the cut-off points used were chosen to create generous circumstances so that if the conjecture failed to make accurate predictions it would be more clear that the conjecture is not worthwhile, if the EIS creates ideal conditions for the conjecture and it still fails, it is more easy to discard the conjecture.

7.5 Summary

In this study, the *awareness* and *flexibility* conjecture and present EI theory were tested using a similar design to that of chapter 6. Based on the conjecture, it was predicted that higher EI would be associated with increased affect regulation (i.e., less reactivity), but *only* amongst low-EI individuals. High-EI individuals were expected to regulate affect in an unpredictable manner. These predictions were only tested in the 'ambiguous' experimental condition because it was expected that *awareness* and *flexibility* would play little role in participants' behaviour when they were given an explicit instruction. Present EI theory leads to the expectation that will be a (overall) correlation between EI scores and reactivity to the affect induction (i.e., affect regulation) when participants were explicitly instructed to neutralise their affect. After completing an EI test, participants' skin conductance was recorded while they were presented with a series of images which they rated in terms of affective valence and intensity. The self-report and skin conductance results provided partial support for the conjecture-driven hypotheses and very limited support for the hypotheses driven by current EI theory. These results cast some doubt on present EI theory but suggest that the conjecture is not a perfect alternative. Potential interpretations of these findings were discussed and it was concluded that a replication using a different EI test would help rule out whether or not the results of the previous two studies were due solely to the test used.

Chapter 8: Prompted Affect Regulation and the *Awareness* and *Flexibility* Conjecture – Self-report and Physiological Indices 2

8.1 Introduction

The results of the prior study provided partial support for the *awareness* and *flexibility* conjecture and limited support for present EI theory. The results were similar to the findings of study 6, but it was determined that another replication would help clarify which results are robust across experiments. It also seemed sensible that this replication should make use of a different EI test to ensure that previous results were not observed solely because of the particular EI test used.

The only meaningful difference between this study and the previous study is the use of a different measure to assess EI and this was done for two major reasons. The first reason was to determine if the results of the previous studies were due only to the EI measure used. Secondly, it seemed worthwhile to determine if using a different EI test would strengthen or clarify the findings. Austin (2004) and Farrelly and Austin (in press) used different tests for precisely this purpose.

The logic of this study is identical to that of the previous study. Just as in the previous two studies, this study has been designed to test hypotheses based on either the conjecture or present EI theory using two experimental conditions. The intent is to determine whether or not EI test scores correlate with degree of reactivity to affect induction. Greater reactivity is indicative of less regulation thus this study investigates whether or not EI test scores are related to lab indices of affect regulation. The ambiguous condition has been designed to test the conjecture: given an ambiguous stimulus, it is expected that low-EI individuals will neutralise their affect in line with their EI scores (i.e., there will be a negative association between EI test scores and reactivity to the affect induction) but high-EI individuals will not (i.e., no such association). The clear condition is included to test current EI theory, as it would be expected that EI scores will correlate with affect neutralisation (i.e., a negative association between EI test scores and reactivity to the affect induction) when participants are given explicit instructions to neutralise affect.

As in chapter 7, self-report measures are included as a manipulation check and as an incidental probe of whether the self-report data return different findings than the physiological data. As was noted previously, it is not appropriate to use a single dependent variable to determine if the manipulation was effective *and* test the hypothesis. The predictions and results for the self-report investigation should be taken as an interesting aside only. The main role of the self-report data was to ensure that participants were indeed regulating affect when requested to do so in the clear condition. This frees the physiological data to be used to test the hypotheses.

It may be argued that at this point a different experimental direction would be better than an additional replication of this study. The preceding experiments have not found conclusive evidence in favour or in opposition to the conjecture so perhaps it would be sensible to abandon this line of research. This is probably a entirely sensible viewpoint, but in this case it was felt that thoroughness in one experimental paradigm would be preferable to a superficial survey of numerous research avenues.

There are two sets of hypotheses for this study. Based on the conjecture, it is predicted that in the ambiguous condition, EI scores will correlate with SCR amplitudes amongst low-EI individuals only, and not amongst high-EI individuals. According to current EI theory, it is predicted that there will be an overall correlation between EI scores and SCR amplitude in the clear condition. The incidental predictions for the self-report data differ slightly. Because the self-report data are not 'change' data (i.e., they do not control for baseline affective status) they are more informative as a measure of how much a stimulus impacted an individual. Based on current EI theory it would be expected that in the clear condition, EI would be positively associated with ratings of valence (i.e., higher EI is associated with a more positive impact of the stimulus) and negatively associated with ratings of intensity (i.e., higher EI is associated with less intense impact of the stimulus). This would be expected because higher EI is associated with greater affect regulation abilities and these abilities should insulate higher-EI individuals from the negative impact of stimuli. In the ambiguous condition it is predicted that EI test scores will be positively associated with valence ratings and negatively associated with intensity ratings amongst low EI individuals but there will be no correlation between these variables amongst high-EI individuals.

8.2 Method

8.2.1 Design

The design of this study was identical to the previous study. Participants completed a measure of *trait* EI before taking part in the main experimental routine. During the experiment, self-report ratings of affective state and SCR were recorded while participants viewed images with either negative, neutral, or positive affective weightings in two experimental conditions (ambiguous or clear). Z-transformed SCR amplitudes and self-report ratings of valence and intensity were correlated with scores on an EI measure in a counterbalanced within-subjects design.

8.2.2 Participants

Thirty seven (twenty-five female, twelve male) students of Edinburgh University took part in this study for pay. Their ages ranged from nineteen to thirty three (M = 23.7, SD = 3.5) Participants were recruited using and advert on the student employment website and through personal contacts. They were paid £2.50 for their participation.

8.2.3 Measures

The 41-item adaptation of the EIS (Austin, Saklofske, Huang, & McKenney, 2004) was used in this study. Items are responded to on a 7-point Likert scale anchored at 'strongly disagree' and 'strongly agree' in a self-report paradigm similar to that of personality and other *trait* EI measures. The reliability of the responses was acceptable in this sample ($\alpha = .73$).

8.2.4 Apparatus

The same experimental materials were used in this study as in the previous study. A touch sensitive monitor was used to present all stimuli and also served as an input device for participants' self-report ratings of affective reaction. The computer program used to present the stimuli and record the indices of arousal (SCR and self-report) was identical to that of the previous study and was designed by Paul Stevens. As in chapters 5-7, skin conductance was recorded at approximately 40 Hz and all data were z-transformed.

8.2.5 Procedure

The procedure of this study was also identical to the previous study, with the exception that a different EI measure was used. Participants completed the EI measure before being attached to the SCR electrodes. After the experiment was explained to them, participants took part in either the ambiguous or clear condition in a counter-balanced order. In both experimental conditions, participants viewed images with positive, negative, and neutral affective loading and rated them for their valence (positive/negative) and intensity (mild/intense) while SCR was recorded for five seconds. All images were presented for five seconds and a random-length pause was included between pictures to prevent expectation effects.

8.3 Results

Table 8.1 displays descriptive statistics for the self-report indices of affect. These data were used primarily as a manipulation check but the incidental hypotheses are tested below as well. The data indicate that the manipulation was less effective than in the prior study. There are significant differences between the valence ratings of positive and negative images between experimental conditions ($t(37) = 4.12$, $p < .001$; $t(37) = -5.33$, $p < .001$, respectively), one of which indicates that ratings got *more* extreme when participants were explicitly asked to neutralise their mood – the opposite of what was expected. There were no significant differences in intensity except in the neutral pictures conditions ($t(37) = 6.17$, $p < .001$).

Table 8.1 Descriptive statistics for self-report indices of affect

		Affective Condition		
Task Condition		Positive	Negative	Neutral
Ambiguous	Valence	3.7 *(1.0)*	4.0 *(1.0)*	3.8 *(1.7)*
	Intensity	4.2 *(1.6)*	3.9 *(1.5)*	4.1 *(0.9)*
Clear	Valence	4.7 *(.7)*	3.2 *(.85)*	4.0 *(.5)*
	Intensity	3.9 *(1.3)*	3.7 *(1.4)*	2.6 *(1.4)*

Note: Values in parentheses are the standard deviations for the condition.

To test the conjecture-driven hypothesis, sub-groups were formed based on a cut-off score approximately equal to the mean ($M = 200$) and EI test scores were correlated with the self-report ratings in the ambiguous condition. In the ambiguous condition, the low-EI sub-group ($N = 19$) showed a positive correlation between EI scores and ratings of the positive images ($\rho = .60$, $p < .01$) only. There was a relationship between EI test scores and self-report ratings in the overall sample which seems to support the prediction based on current EI theory. The association between EI and self-report ratings in the ambiguous condition in the high-EI group ($\rho = .55$, $p < .01$) was unexpected and seems to disconfirm the conjecture. Again, the results of these self-report hypothesis tests should be read as a reference only as the self-report indices of this study were included to serve as a manipulation check only. Table 8.2 below displays the correlations between EI test scores and SCR amplitude in the different experimental and EI sub-group conditions.

Table 8.2. Correlations between EI test scores and SCR amplitude in the different experimental and EI sub-group conditions.

EI Subgroup	Self-Report Rating of Affective State							
	Intensity				Valence			
	Am.P	Am.N	C.P	C.N	Am.P	Am.N	C.P.	C.N.
EI test score (Overall, $N = 37$)	0.31	0.05	0.17	0.16	0.22	**-0.37**	0.05	0.12
EI test score (high, $N = 18$)	**0.55**	-0.13	0.37	0.32	-0.16	-0.16	0.05	0.15
EI test score (low, $N = 19$)	**0.60**	0.06	0.16	0.06	0.09	0.33	0.22	0.04

Note: Am. P refers to the ambiguous positive condition, Am. N to the ambiguous negative condition, C. P. to the clear positive condition, and C. N. to the clear negative condition. Different rows display correlations in either the overall sample, the high-scoring group, or the low-scoring group. Bold numbers are significant at the $p < .05$ level.

Table 8.3 displays the descriptive statistics for the physiological indices of affect in the various experimental and affective conditions. As in the previous study, there was a relaxation trend in all conditions and as a result, some conditions initially appear to have resulted in a negative SCR amplitudes. These descriptive statistics are averages of individuals' SCR amplitudes after z-transformation and as such are expressed in terms of standard deviations from the individual's overall session mean. They are provided solely for reference. Group differences in SCR are not of interest here, rather it is the relationship between EI test scores and individual changes in arousal which is of concern.

Table 8.3. Descriptive statistics for physiological measures of affect

Task Condition		Affective Condition		
		Positive	Negative	Neutral
Ambiguous	Mean	.04	-.07	.09
	S.D.	.34	.24	.28
Clear	Mean	.07	.06	-.06
	S.D.	.30	.27	.32

Note: All values are expressed in terms of standard deviations off the mean (σ).

After the sample was divided into high- and low-scoring sub-groups, correlations between EI scores and SCR amplitude in the ambiguous condition were tested. When divided at the mean (M =200, N = 18 in the high, 19 in the low), the moderate positive correlation between EI score and SCR amplitude amongst the low-EI individuals in the negative ambiguous condition approached significance ($\rho = .33$, $p < .08$) and there was no correlation between EI scores and SCR amplitude in the negative condition nor were there any correlations between EI score and SCR amplitude in the high-EI group. These findings fail to support the conjecture-driven hypotheses. Figures 8.1 and 8.2 illustrate this slightly weaker 'neutralisation' effect which is present for low-EI individuals but which is not present in the entire sample. This indicates that the hypothesized relationship between EI test score and reactivity to affect induction exists for low-EI individuals but not for high-EI individuals. The predicted overall correlation between EI test scores and SCR amplitude was not observed and thus it is not possible to reject the null hypothesis in favour of the hypothesis driven by current EI theory. All of the correlations between EI test scores and SCR amplitude are presented below in Table 8.4.

Table 8.4. Correlations between EI test scores and SCR amplitude in different experimental conditions and EI sub-groups.

	Z-Transformed SCR Amplitude			
EI Subgroup	Amb. P	Amb. N	Clear P.	Clear N.
EI test score (Overall, N = 37)	-0.14	-0.09	-0.06	0.11
EI test score (high, N = 18)	0.04	-0.27	-0.23	0.16
EI test score (low, N = 19)	-0.01	-0.33	0.26	-0.22

Note: Amb. P refers to the ambiguous positive condition, Amb. N to the ambiguous negative condition, Clear P. to the clear positive condition, and Clear N to the clear negative condition. Different rows display correlations in either the overall sample, the high-scoring group, or the low-scoring group. Bold numbers are significant at the $p < .05$ level.

Figure 8.1 The relationship between EI scores and SCR amplitude in response to negative stimuli in the ambiguous condition – low-EI participants only

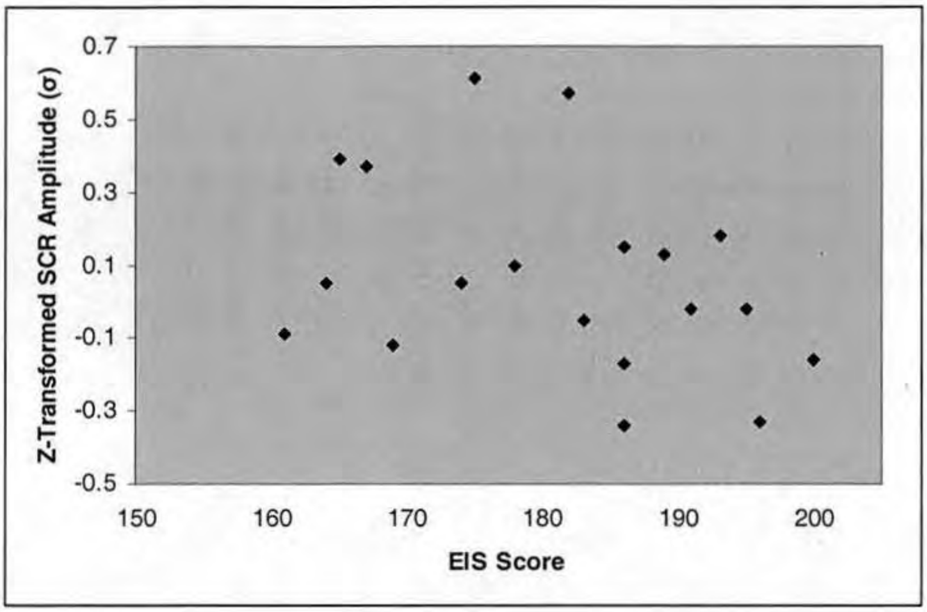
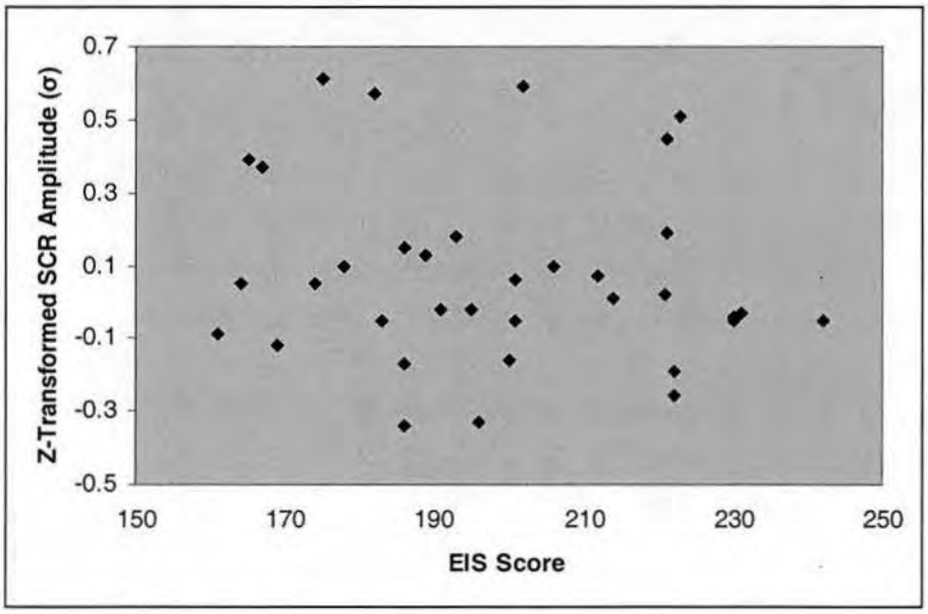


Figure 8.2 The relationship between EI scores and SCR amplitude in response to negative stimuli in the ambiguous condition – entire sample.



Again, these figures illustrate the 'neutralisation' effect that disappears beyond a certain point – people with middling EI are not affected by the stimuli whilst very low-EI individuals show less reactivity. As in the previous studies, high-EI individuals' affective reactions do not correlate with their EI scores.

8.4 Discussion

The self report results showed that amongst low-EI individuals, EI scores were correlated with ratings of affective intensity in the ambiguous condition, but only for positive images, suggesting that higher-EI was associated with greater intensity of positive affect for these images ($\rho = .60$, $p < .01$). These findings fail to support the conjecture. The lack of correlations between EI scores and self-report indices of affect in the clear condition indicate that the null hypothesis could not be rejected in favour of the hypothesis driven by current EI theory. Again it should be clarified that these self-report hypotheses are included and tested as a reference only because it is not exactly appropriate to use the subjective measures as a manipulation check and a dependent variable. They merely serve as a superficial probe of whether or not the relationship between EI test scores and reactivity to affect induction depends on the measure used.

The physiological data revealed a marginally significant correlation between EI scores and SCR amplitude, amongst low-EI individuals, in the positive pictures condition (.33, $p < .07$), but no correlations between EI scores and SCR amplitude for the high-EI individuals. There were no overall correlations between EI scores and SCR amplitude in the clear condition. Thus there appears to be partial support for the *awareness* and *flexibility* conjecture but the null hypothesis must be retained instead of the hypothesis based on current EI theory.

In light of this replication which used a different EI measure from the previous two studies, it would seem that there is some degree of very limited support for the *awareness* and *flexibility* conjecture. This topic will be discussed at length in the next chapter, but it suffices to say here that altering the EI measure does not appreciably impact the results obtained from this series of studies.

The sample in this study was slightly larger than in the previous study, but as in all previous studies, it certainly could have been larger. Psychophysiological measures are characterised by large individual differences, so it is possible that a larger sample would reveal clearer findings. Sampling issues will be discussed in more detail in the general discussion.

The experimental manipulations in this study seemed less successful than in previous studies and it is difficult to understand why. The neutral images seemed to elicit the expected reactions, but the positive and negative images seemed to function differently in different experimental conditions, which themselves seemed to have the opposite effect from what was expected. For this reason, this experiment should probably carry slightly less weight than the preceding two studies, but despite this flaw, there was still some support for the conjecture.

At this point, it seems unlikely that a future replication would shed any additional light on the conjecture or on present EI theory. Different EI tests could be used, as could different stimuli or any number of other small aspects of the experiment. However, these would be unlikely to drastically alter the findings that have emerged from chapters 5-8. As such, these results are probably sufficient to draw some final conclusions about the conjecture and current EI theory.

8.5 Summary

This chapter was a replication of the previous study with a different EI test. Participants completed the EIS and then took part in an experimental session otherwise identical to the previous study. No conclusive support was found for the conjecture as the correlation between EI scores and SCR amplitude in response to positive pictures in the ambiguous condition only approached significance, but only for low-EI individuals. The lack of any overall correlations between EI score and

SCR amplitude suggests that the null hypothesis could not be rejected in favour of the hypothesis based on current EI theory.

Chapter 9: EI and Flexibility in Coping

The purpose of this study was to examine some predictions made based on the conjecture with a method that differed from the approach used in chapters 5 through 8. This study was designed to methodologically different but to test similar predictions as were tested in those studies, but this study also includes some novel hypotheses which are appropriate to these data.

9.1.1 Coping Styles

As was mentioned in chapter 1, EI has long been discussed alongside coping. The similarities between the two constructs are quite apparent as both concern how individuals manage the emotional content of their lives. Coping is also discussed frequently in popular psychology books but it is a more complex construct than is apparent from lay conversation.

The word 'coping' actually refers to a number of emotion-related processes or 'styles'. There has been some disagreement and refinement of what precisely these styles are, but it is fairly well agreed-upon that there are four major coping styles: rational, emotion-focused, detached, and avoidant (Roger, Jarvis, & Najarian, 1993). These four styles refer (respectively) to an approach characterised by activity and reasoning, introspective consideration, emotional distancing, and denial. This categorisation into four styles is intuitively appealing and has also received psychometric support (Roger et al., 1993, Elklit, 1996). One modern instrument which can be used to measure coping styles is the Coping Styles Questionnaire (CSQ). This measure is based on previous measures which posited only three coping styles but it contains items which measure the fourth style, detached coping.

Responses to the CSQ have been shown to support the four-factor structure which it was designed to measure and have also been shown to support the convergent and divergent validity of the measure (Roger et al., 1993, Elklit, 1996). It seems that the four coping styles are obliquely related to one another although these relationships are weak enough in most cases to suggest a large amount of independence between the coping styles. More specifically, rational coping is moderately (i.e., $r = -.5$) negatively associated with emotion-focused coping, but the relationship between other coping styles are fairly weak and are often statistically nonsignificant. Scores on the CSQ have also been shown to correlate with personality scores. Specifically, it seems that scores on the emotional coping style correlate positively with the ruminative, maladaptive aspects of neuroticism and that scores on the rational coping style correlate with the surgency aspect of extroversion (Roget et al., 1993). These connections between coping styles and personality processes seem to suggest that coping techniques have important ramifications for affective state.

9.1.2 EI, the Conjecture, and Coping

In some of the popular myths about EI, it is argued to be something that protects individuals from all maladaptive emotions or emotional processes. While this is almost certainly overstated, even the most conservative EI theorist would probably argue that there would be some correlation between EI test scores and measures of coping styles (e.g., the CSQ). A *trait* theorist might argue that the dispositions and self-beliefs that comprise EI (e.g., statements such as 'I know how to handle my reactions to bad news) are similar in content to what would be measured on an index of coping styles. An *ability* theorist would argue that because high-EI individuals are more skilled at regulating their emotions, they would (almost by definition) be better at coping.

Although the existence of *some* relationship between EI and coping styles is fairly obvious, it is less clear *which* coping styles would be more likely to be used by high- or low-EI individuals. Rational coping seems very similar to EI in that both are concerned with using cognition to influence emotions. Specifically, rational coping seems similar to the affect regulation facet of EI, so it would seem that there would be a positive association between EI test scores and rational coping scores. Emotion-focused coping also seems like it would be associated with EI. This form of coping involves introspection and reflection so it would seem likely that it would correlate with the intrapersonal facets of EI. However, unlike intrapersonal EI, emotion-focused coping tends to involve dwelling on the negative or unchangeable parts of the stimulus which requires coping (e.g., it's out of my control so I might as well get used to it), so it seems likely that a negative relationship between EI test scores and emotion-focused coping would emerge. Also, it is known that EI is negatively associated with neuroticism and that emotion-focused coping is positively associated with neuroticism, and although this is far from ironclad logic, it seems intuitive that EI would be negatively associated with emotion-focused coping.

The remaining two coping styles are somewhat more difficult to connect to either *trait* or *ability* EI. Coping through avoidance is a maladaptive technique that seems to have little to do with emotions at all. If anything, this style of coping is an active choice to *not* engage with one's affective state. That said, there do seem to be superficial similarities between low-EI individuals (i.e., people who are bad at identifying and handling their emotions) and people who simply avoid emotions where possible. The potential connection between these constructs is explored in this study. Although detached coping is similar to avoidance coping in that it also involves emotional disengagement, it is far more adaptive in that usually is characterised by an attitude of 'although it is out of my control, I'm going to do everything in my power.' Rather than simple avoidance, detached coping involves an effort to avoid the negative aspects of whatever stimulus one is coping with whilst simultaneously exploring what is within one's realm of control. The classic example for this sort of coping is a terminally ill patient. This sort of technique does not seem nearly as similar to EI as rational or emotion-focused coping do. It also does not fit snugly into the personality traits which are normally associated with EI (e.g., neuroticism and extroversion).

The relationship between EI and coping was testing by Saklofske, Austin, Galloway, and Davidson (2007). It was found that *trait* EI was negatively associated with emotion-focused coping and positive associated with rational coping, just as would probably be expected according to either EI theory. However, they did not examine the potential relationship between EI and the other two forms of coping.

The detached and avoidant coping styles provide an excellent testing ground for the conjecture. In chapters 6 through 8, the ambiguous condition was the critical condition for testing the hypotheses based on the conjecture. As was noted in chapter 6, stimulus ambiguity is essential for the conjecture to be observed, for if the response was clearly elicited, there would be no room for *flexibility*. The detached and avoidant coping styles provide exactly the ambiguity required to test the conjecture. As was just noted, avoidant coping is generally considered maladaptive and detached coping is generally considered somewhat adaptive. However, these two styles seem less clearly adaptive or maladaptive than rational or emotion-focused coping. That is, it is almost always a good idea to be proactive and assertive and almost always a bad idea to ruminate on a negative situation. The utility of the other two styles of coping is more situational. Avoidance coping could well be a viable 'second choice' of techniques when rational coping fails and in some circumstances (e.g., the death of a loved one) it may be the only sensible option. It is more difficult to be prescriptive with this style of coping or with detached coping.

Thus these two styles of coping provide for an examination of *flexibility* which is similar to chapters 6 through 8. The conjecture leads to the prediction that high-EI individuals will display more variability in their employment of these coping styles because of their increased *flexibility* in response to the demands of the situation. Low-EI individuals on the other hand would be expected to be less variable in how they use these types of coping – they will either generally avoid it or generally utilise it. These are similar predictions to those made about the SCR data in chapters 6 through 8.

This study will also test the conjecture using another type of analysis. In chapters 5 through 8, *flexibility* was operationalised as variability. This is a perfectly sensible operational definition, given that flexibility by definition entails change, dynamics, or variability. However, the nature of the conjecture and the nature of these different coping techniques also allows an additional operational definition of *flexibility* and *awareness*.

The coping style data examined in this study show how the incremental value of the conjecture to EI theory. It is already known that EI is related to rational and emotion-focused coping. This is fairly straightforward, but as was already discussed, the relationship between EI and the other types of coping is not as clear if only existing EI theory is considered. The conjecture can help clarify how scores on measures of these other two styles of coping may relate to EI test scores. Specifically, the conjecture leads to the prediction that because of their increased *flexibility*, high-EI individuals would be more likely to use these two types of coping *in situations where the other types of coping are not.* In other words, the conjecture makes predictions about which styles of coping high-EI individuals use *over and above* rational and emotion-focused coping. To use statistical language, the conjecture makes a prediction about how EI test scores relate to detached and avoidance coping when the variance from rational and emotion-focused coping is partialled out. That is, current EI theory informs us about how EI may relate to rational and emotion-focused coping but is less useful for understanding the relationship between EI and avoidant or detached coping. The conjecture 'fills the gap' by make a prediction about how EI may relate to these remaining two styles of coping after controlling for the styles of coping we can already make predictions about. Thus the conjecture-driven hypotheses for this study are a perfect example of how the conjecture provides an incremental theoretical step beyond existing EI theory.

Of course, it should be noted that the regression model used to test the conjecture here is not a direct measure of whether or not individual people do in fact switch coping styles to suit situations. It is not possible to measure strategic use of coping styles with either the CSQ or the EIS as neither. However, what this test will reveal is whether or not individuals *as a group* tend to use particular coping techniques in line with their EI scores, after the primary coping styles (rational and emotion-focused) are controlled for.

9.1.3 Specific theoretical issues

The preceding studies have all examined high- and low-scoring groups but the methods used to create these groups has not been discussed in any detail. This is an unacceptable omission given the importance of these groups.

To assess the conjecture in chapters 5 through 8, the samples were split into high and low-scoring groups at some point on or near the mean score. As was discussed in previous chapters, the cut-off point used to divide groups is indeed an arbitrary decision and in the preceding studies, the most generous cut-off score has been used. Obviously this process of using the cut-off score which yielded the strongest effects provided an artificial 'boost' to the results in favour of the conjecture and all of the results argued to support the conjecture must be interpreted with this in mind. Equally obviously, no firm conclusions could be based on such a practice given its arbitrariness. However, the reason for this ultimately dubious practice was to allow for the best possible circumstances for the conjecture so that if the conjecture failed to make accurate predictions *in spite* of this methodological generosity, it would be even easier to discard the conjecture as a theoretical possibility. The intent was to use the most generous methods possible and then become more stringent if the conjecture was well supported.

However, there are other options for creating sub-groups. A tempting option would be to use the mean or median value from a normative population. There are two major problems with this method, the first is that university students are not the normal population so assuming the mean from the normal population applies to them would be erroneous. This 'canned mean' problem is more important than it may seem. The accuracy of the population mean as a model for overall performance of any sample depends entirely on whether or not the sample comes from that population. It is patently obvious that high- and low-scoring autistic individuals would not split according to the mean IQ of the normal population and the situation in EI research on university students is not entirely dissimilar. The most accurate mean model is the mean drawn from a given sample. The second issue is that many EI tests have not in fact been normed on the general population so it is not possible to know what the population mean is.

Still, some sort of external justification for using a sample mean as a cut-off would strengthen the methodology of this study. Although a mean or median split is entirely acceptable, using such procedures make it impossible to know if the participants are 'high-EI' or 'low-EI' in any objective sense because the sub-groups have been constructed based on a sample-bound value. Two simple methods can be used to address this issue. One would be to examine past research on the EI measure used on this study to determine how the mean value found in this study compares to other samples. Another option would be to use smaller extreme groups to test the hypotheses rather than make larger 'high' or 'low' groups. The latter approach is used in this study because if the extreme groups are created sensibly (i.e., using the top and bottom 10% of scorers) then it is almost certain that the participants in these extreme groups will be objectively 'high-EI' or 'low-EI'. The extreme groups approach is also used for expediency's sake (a meta-analysis of every available dataset for the EIS would be highly laborious and would only raise additional issues such as how a 'grand university student average' would be calculated) and because this approach has a historical precedent for being used to examine the relationship between test scores and theoretically-related variables. The use of extreme groups will allow more certainty that the high- and low-EI groups do in fact represent distinct sub-populations.

9.1.3 Hypotheses

Based on the conjecture it is predicted that there will be a negative association between EI scores and avoidance coping amongst the low-EI individuals but that this relationship will not exist amongst high-EI individuals. A positive association between EI scores and detached coping is predicted amongst the low-EI individuals but it is also predicted that there will be no such association amongst the high-EI individuals. It is also predicted that when the variance attributable to rational and emotion-focused coping is partialled out, there will be a positive association between EI test scores and scores on detached and avoidance coping, due to the fact that they are both probably adaptive and 'emotionally intelligent' if rational coping can not be used for some reason. These coping styles will be the focus of this study, but as was previously mentioned, emotion-focused and rational coping are only used as covariants.

9.2 Method

9.2.1 Participants and Design

An opportunity sample of 364 participants (258 females, 104 males, two non-respondants) took part in this study. The participants were all university students in Canada. The mean age of these participants was 24 (SD 6.1). The participants completed the unadapted Schutte et al., (1998) EIS, a personality measure, the CSQ, and also provided some self-reports of their exercise and lifestyle behaviours. Responses to the EI and coping styles measures are analysed using a correlational design.

9.2.2 Measures

33-item EIS. The Schutte et al., (1998) EIS was used to measure *trait* EI. Additional information about this measure is presented in chapter 4. In this sample the internal consistency reliability of this measure was high ($\alpha = .90$).

CSQ. The shortened version (Elklit, 1996) of the Coping Styles Questionnaire (Rogers et al., 1993) is a self-report measure which uses Likert scales to test coping behaviours which assess four types of coping style: rational, emotion-focused, detached, and avoidant. This measure has 37 items. In this sample the internal consistency reliability of rational coping and emotion-focused coping were acceptable ($\alpha = .82, .76$) detached coping was below acceptable levels ($\alpha = .56$) and avoidance coping bordered on acceptable ($\alpha = .65$).

The additional measures used are described in more detail in Saklofske et al. (2007) and are not explicated here because these data were not analysed in the present study.

9.2.3 Procedure

The data presented in Saklofske et al.(2007) were analysed for this study. These data were collected by the primary researchers involved in that study. Prospective participants were told that the study was concerned with personal factors and health behaviours.

To create sub-groups to test the first set of hypotheses, participants in the top and bottom 10% of scorers were assigned to the high-EI and low-EI groups, respectively.

9.3 Results

Table 9.1 presents the descriptive statistics for the EI measure and the four coping style measures. These values are for reference only and have no bearing on the hypothesis.

Table 9.1 Descriptive statistics for EI and coping styles measures

Measure	Mean	SD
EI	123	14
Rcope	29	5
Ecope	17	5
Dcope	13	2
Acope	22	4

Note: EI refers to scores on the EIS, Rcope to rational coping style, Ecope to emotion-focused coping, Dcope to detached coping, and Acope to avoidant coping.

The first set of hypotheses was examined in a similar manner to chapters 5 through 8 with the important exception that extreme groups were used. Those participants scoring in the 10% or less on the EI measure formed the low-EI group. This group was made of 35 participants ($M = 95$, $SD = 14$). The top 10% of scorers formed the high-EI group. This group was comprised of 41 individuals with a mean score of 142 ($SD = 5$). These sub-groups were used to examine the conjecture-driven hypothesis, although it would seem that there is *prima facie* evidence against the conjecture in the finding that this high-EI extreme group had much more consistent EI scores than their low-EI counterparts.

When only low-EI individuals were examined, there were no associations between EI test scores and either of the coping styles (all p's > .1). The null predictions amongst the high-EI individuals were supported by the finding that EI was nonsignificantly (all p' s > 1) related to avoidant and detached coping amongst the high-scoring sub-group. These findings do not justify a rejection of the null hypothesis and thus the first two predictions driven by the conjecture were not supported.

Two hierarchical linear regressions were carried out to test the second group of hypotheses driven by the conjecture. There were two blocks entered into both of these analyses in order to test the hypothesis by determining if EI scores predicted detached and avoidant coping after the variance attributable to rational and emotion-focused coping was partialled out. In both regressions, scores on rational and emotion-focused coping were entered as the first block and EI was entered as the second block. This is because it is already known that the coping styles are interrelated and because the hypothesis concerned whether or not EI predicted variance in detached and avoidant coping after controlling for rational and emotion-focused coping.

The analysis was then carried out to determine if including EI test scores in the regression model predicted a significantly larger amount of variance in the outcome measures (i.e., detached and avoidance coping) than rational and emotion-focused coping below. In other words, the incremental predictive utility of EI test scores for detached and avoidant coping styles was examined. A breakdown of the blocks (i.e., models) and their predictive utility is presented in Table 9.2 and 9.3.

Table 9.2 Incremental predictive value of EI for detached coping

Model	Variables	*r*	R^2 change	Sig
1	Rcope, Ecope	0.447	0.2	**0.01**
2	Rcope, Ecope, EI	0.454	0.006	0.14

Note: Rcope refers to rational coping, Ecope refers to emotion-focused coping, EI refers to the EIS, Sig refers to the significance of the R^2 change between models.

Table 9.3 Incremental predictive value of EI for avoidant coping

Model	Variables	*r*	R^2 change	Sig
1	Rcope, Ecope	0.338	0.109	**0.01**
2	Rcope, Ecope, EI	0.338	0	0.78

Note: Rcope refers to rational coping, Ecope refers to emotion-focused coping, EI refers to the EIS, Sig refers to the significance of the R^2 change between models.

As is clear from Tables 9.2 and 9.3, the inclusion of EI as a second model did not account for any additional variance in detached or avoidant coping scores. Thus the second set of hypotheses based on the conjecture can not be supported and the results from this study fail to support any of the predictions made based on the conjecture.

9.4 Discussion

The results of this study failed to confirm any of the predictions made based on the conjecture. The first set of hypotheses predicted that EI would relate to avoidance and detached coping in a manner similar to the relationship predicted between EI and SCR amplitude in chapters 5 through 8. Namely, it was expected that there would be a correlation between EI test scores and avoidant and detached coping amongst the low-EI individuals but that this relationship would not be present amongst the high-EI individuals. The predicted relationship amongst the low-EI individuals was not found (all p's > .1) and although the predicted nulls were found, it is probably inaccurate to argue that these nulls support the conjecture. This issue of 'nulls as support' is returned to in the discussion.

The second set of hypotheses predicted that due to their increased *flexibility,* high-EI individuals would be more likely to use detached and avoidant coping in situations where rational or emotion-focused coping failed or was otherwise inappropriate. In other words, it was predicted that once the variance attributable to ration or emotion-focused coping was controlled for, there would be a significant relationship between EI test scores and avoidant and detached coping. This hypothesis was not supported as EI failed to predict a significant amount of variance in these types of coping when rational and emotion-focused coping were partialled out of the model.

This study provides a rather conclusive end to this series of studies of the *awareness* and *flexibility* conjecture. The coping styles that were of chief interest in this study were expected to be excellent examples of the sort of variable that the conjecture would be most useful for. That is, the utility of these coping styles is highly situational and ambiguous (just like unprompted affect regulation) and thus should have been exactly the sort of emotion-related process that the conjecture would have been of great use in predicting. This expectation was not supported by the data.

It also seems that even when an extreme-groups paradigm is used, it is not possible to find clear support for the conjecture. The sub-groups formed in this study by using the top and bottom 10% of scorers were highly different from one another (M = 142, 95, respectively) but it was not possible to find any evidence for the conjecture-driven hypothesis using this more refined technique of sub-group creation.

The internal consistency reliability of the detached and avoidant coping measures was less than ideal in this study and could have contributed to the lack of support for the conjecture. However, this sort of defence seems rather stretched. First of all, it was only detached coping which had notably low internal consistency reliability and this was almost certainly due to a relatively small number of items (only 7). Secondly, the possibly dubious nature of the coping style measure would be of far more concern if the conjecture *had* been supported than, as was found here, if the conjecture was not supported. To put it simply, the conjecture has received support that is limited at best and, frankly put, it is probably not the methodology which needs refinement. The results of this study clearly fail to support any of the hypotheses made based on the conjecture and as such the only responsible conclusion to make is that the conjecture is probably not a useful addition to EI theory. It is with this in mind that general issues are discussed.

Chapter 10: General Discussion 1 – Relevance and Interpretations of Findings to Present EI theory and the *awareness* and *flexibility* Conjecture

The purpose of the studies in this thesis was to test hypotheses based on either current EI theory or the *awareness* and *flexibility* conjecture. A psychometric investigation, a series of four lab experiments, and a simple regression analysis were carried out for this purpose. The results from these studies generally failed to support the conjecture. The findings from these studies also failed to provide conclusive support for the hypotheses driven by current EI theory. This chapter examines the findings in more detail, their relevance to their respective theories, and also explores other reasons these findings may have occurred.

10.1 Review of Findings

10.1.1 Findings Fail to Support Predictions based on Current Theory

The studies in this thesis have largely failed to support hypotheses based on current EI theory. The factor structure of EI tests fails a basic psychometric assumption, there was no overall correlation between EI scores and SCL change from baseline, and studies 6-8 all found no correlation between EI scores and SCR amplitude in the clear conditions. Chapter 9 did not examine any hypotheses based on current theory because the data analysed in that study have already been used to show that EI test scores correlate with coping styles. Although these findings were discussed in the various chapters, more detailed commentary about the relevance of these findings seems necessary.

In chapter 4, a fundamental assumption underpinning EI theory – the assumption of factorial invariance – was shown to be false for EI tests. Factor solutions were not consistent between sub-groups. Rather, it seems that high-EI and low-EI individuals' responses to EI questions have notably dissimilar factor loadings such that the factors that underpin low- and high-scorers' responses are not constructed of the same items. This finding would be analogous to finding that the fluid and crystallised ability factors extracted from IQ test results differed according to IQ level such that a verbal item only loaded on crystallised ability in one sub-group. If such a finding occurred in IQ research, it would violate what is known about intelligence, and now a similar dilemma faces EI researchers, for this finding not only violates the assumptions behind EI theory but it also has important practical ramifications. This is why differentiation in IQ is so hotly contested.

The finding that EI tests do not measure high- and low-scorers on the same dimension presents both theoretical and practical problems to the EI testing enterprise. In addition to the IQ analogy used above, this finding is also analogous to discovering a ruler that did not measure tall and short people on the same dimension (i.e., height). Not only would most people be very sceptical about the utility of such a ruler, but we might also wonder which of the two dimensions was correct. It is hard to know what exactly we should make of EI tests which do not measure all test-takers on the same dimension, and it is hard to explain away these findings. Practical difficulties associated with factorial variance also present themselves. It is well known that differences in the explanatory power of the first principal component are associated with differences in reliability (Hartmann & Teasdale, 2005). A *post-hoc* examination of the data in chapter 4 revealed that there were substantial differences in the alpha estimates (sometimes up to a .15) between high- and low-scoring groups, so it would appear that differences in factor structure are associated with important practical differences. Thus it would seem from the results of chapter 4 that EI tests are beset with considerable theoretical and practical difficulties as a result of their factorial variability.

The results of Chapter 5 also contraindicated the hypothesis made according to EI theory. It was predicted in this study that there would be a simple negative correlation between EI score and reactivity to affect induction (SCL change from baseline). This connection was taken to be indicative of a relationship between EI test scores and a lab-based instantiation of affect regulation: smaller reactivity to affect induction is indicative of affect regulation. However, there was no discernable correlation between EI scores and SCL changes in baseline and this seems to suggest that higher EI is not associated with greater abilities at regulating affect at all. Were there a correlation in either direction it would be possible to integrate the finding into current EI theory (i.e., EI relates to a 'control' or 'sensitivity' effect), but the lack of any relationship (similar to the findings of Ciarrochi, Chan, & Caputi, 2000) is difficult to understand in light of current EI theory.

The studies in chapters 6 through 8 also failed to support the hypotheses driven by current EI theory. These studies were very similar to one another as each contained an experimental condition in which all participants were explicitly instructed to neutralise any emotions they felt while viewing the stimuli. All participants seemed to understand the instructions, yet there simply were not any overall (i.e., the entire sample) correlations between EI scores and SCR amplitudes in any of the clear task conditions in any of the experiments. It was not possible to reject the null hypothesis in favour of the hypotheses drawn from current EI theory, save for one exception in the self-report data. Of course, the self-report hypotheses were meant to be incidental only, so this single significant finding is less convincing than the nulls in the SCR data are.

These failures to reject the null hypothesis are surprising. As was noted in the introduction, there have been a few attempts to investigate the link between EI test scores and lab-based 'mood regulation' tasks. Petrides and Furnham (2003) and Ciarrochi et al (2000) found connections between EI and affect regulation but the direction of the correlations differed, possibly due to different methodology: in the Ciarrochi et al. (2000) study, participants were asked to perform a judgment task which would transparently require a 'clear mind' and thus might well have prompted them to regulate affect; such a prompt was not present in the Petrides and Furnham (2003) study. Although the two studies found that EI related to affective reactivity in different ways, the reliable finding in both was that EI has *some* effect on reactivity to affect induction. Chapters 6 through 8 essentially re-created the two experimental situations represented by these two studies: a condition in which participants were explicitly prompted to regulate affect and a condition in which they were not. These studies should have detected any link between EI test scores and reactivity to affect induction (i.e., affect regulation). A connection between EI level and affect regulation was only very rarely found in these studies, regardless of which EI test was used.

Thus it seems that current EI theory receives hardly any support from the results of these studies. The null hypothesis could be rejected in favour of the experimental hypotheses only twice. At face value, the most sensible conclusion seems to be that the results of these studies fail to support EI theory, but potential mitigating factors are discussed below.

10.1.2 The conjecture is not supported

The data from these studies provided virtually no support for the conjecture. Although some hypotheses were supported, at least as many were not. In some studies, it was possible to reject all of the null predictions in favour of the conjecture-driven hypotheses, but in other studies it was only possible to do so in some cases. A review of the findings of these studies would help highlight the few areas in which this conjecture succeeds.

Unlike chapters 6 through 8, chapter 4 was designed such that the hypotheses were mutually exclusive – either EI tests are factorially invariant or they are not – and the results from that study support the conjecture over current EI theory. The results revealed evidence of factorial variance in EI tests and although this is problematic for current EI theory it fits well into the conjecture. According to the conjecture, ambiguous stimuli such as EI tests (in which a single item often represents a multitude of situations and 'correct' responses) are expected to elicit predictable responses from low-EI individuals but unpredictable responses from high-EI individuals due to the high-EI individuals' greater *awareness* of the variegated solutions possible to the ambiguous item and *flexibility* in employing these various solutions. In short, according to the conjecture it is expected that high- and low-scoring individuals will answer EI tests in a different manner because EI functions in a different manner for them. As was explained in chapter 4, it is possible to rule out any future discussion about tests results because of these psychometric flaws, but such a standpoint rules out a great deal of potentially meaningful discussion, especially if there are no un-flawed EI tests. Of course, the ideal situation from the perspective of the conjecture would be EI tests which were factorially invariant and were also capable of measuring *awareness* and *flexibility*. However, it is only possible to use the tests which already exist, so it may be more sensible to withhold judgment on these 'flaws' and instead take the test results at face value and discuss what theoretical importance factorial invariance might have. If this theoretical exploration is allowed, chapter 4 seems to provide fairly robust evidence that responses to EI tests are more in line with what would be predicted according to the conjecture than what would be predicted according to current EI theory.

Chapters 5 through 8 provide very limited support for the conjecture. The hypotheses made in these chapters were all similar and it was possible to reject the null hypothesis in favour of some of them with the physiological and (incidental) self-report results. The correlations between EI and reactivity to the mood inductions varied from study to study (e.g., positive images rather than negative images). Overall, about one half of the correlations predicted according to this theory were found in the physiological data but the incidental self-report data were less positive. This is hardly unequivocal support but it is also indicative that there is *some* sort of limited effect here.

It would be tempting to claim that the nulls that were predicted and found in these studies support the conjecture, but this must be examined closely. Although the conjecture predicts that high-EI individuals will not regulate affect in a predictable manner (i.e., there will be no correlation between EI test scores and reactivity to the induction due to their increased *awareness* and *flexibility*) and although these nulls were indeed found, it may be very difficult to know that these nulls are indicative of the operation of *awareness* and *flexibility* and not some other mechanism. Studies are normally designed to create a state of affairs that should a positive finding be obtained, there is no way to rule out the desired explanation. This certainty is elusive with nulls. It could easily be claimed that any nulls found are simply a result of test unreliability, methodological flaws, or any number of competing explanations and thus that the conjecture does *not* receive any support from these nulls. However, such an argument seems inconsistent. It is not clear why any potential issues with these studies would cause the null findings but not play a role in the positive findings (e.g., if the nulls are due to reliability issues, then it must be asked how the positive findings occurred in spite of the issue). Simply put, if doubt is cast on the nulls, then it must be cast on the positive findings as well. Presumably there might be some flaws that apply to nulls (e.g., because they are non-effects it is probably useless to ask how such a non-effect occurred) and not to positive findings, and any objection to the use of nulls as support for the conjecture ought to be based on one of these unique flaws.

That said, the issue is less about what caused the null findings and more about the logic behind using a null finding to support a theory. Simply put, it is indeed possible to obtain null findings due to the operation of some mechanism but it is also possible to obtain those findings due to chance, poor methods, or any of the reasons mentioned above. It is not possible to discern why a lack of any relationship was obtained, so if a theory is built around null findings it could very well be built around an actual mechanism, but it could also easily be based on chance, poor methods, etc. It may be possible to accurately conclude that a null finding supports a theory, but it is impossible to be certain and is just as likely that the null findings occurred due to some other factor(s). This would result in a theory built entirely on these other factor(s), not the mechanisms proposed in the theory. In defence of these studies, only results which reject the null hypothesis in favour of a positive prediction are argued to support the conjecture. The methods used in this thesis should probably also be given some credit because the experiments were designed conscientiously and it is hard to conceive of an issue that would affect *only* those comparisons which resulted in null findings. Still, it is probably more conservative to base any conclusions about the conjecture solely on the positive findings and regard the nulls as interesting but unconvincing.

This thesis does not seem to support the conjecture. There does seems to be *some* kind of effect which occurs in accordance with the conjecture but the strength and reliability of this effect is debatable and thus so is the veracity of the conjecture. Few of the positive predictions were supported and predicting nulls is of little use for developing a theory. That said, there seems to be as much support for the conjecture in these theories as there was for EI theory in Austin's (2004) and Farrelly and Austin (in press) studies of EI and lab-based inspection-time tasks. If the accurate null predictions are accepted as evidence then these studies support the conjecture considerably more strongly. It seems the most responsible conclusion to make is that these studies show that neither the conjecture nor current EI theory make accurate predictions about how EI might relate to affect regulation, at least in terms of skin conductance.

10.2 Possible other Explanations for these Findings

Of course it is possible that methodology alone is responsible for these findings. Much of this thesis was concerned with a very specific operational definition of a specific facet of EI theory and it is certainly possible that other experimental paradigms might find evidence in favour of current EI theory. With the exception of chapters 4 and 9, the studies only investigated the physiological and self-report indices of affect in response to an affect manipulation and the lack of any correlation between these indices and EI scores (in overall samples) does not necessarily mean that other indices would not correlate with EI test scores. Of course, chapter 9 partially mitigates this limitation.

It is possible to question the suitability of physiological indicators into account and it is possible that if different indices of affect were used the results would have been markedly different. As was noted in the introduction to this thesis, physiological methods are characterised by massive individual differences and it could be argued that they are therefore inaccurate and these results are unconvincing. It might also be argued that because emotions are first of all a personal, subjective experience rather than change in SCR amplitude, the physiological indicators are irrelevant. Although individual differences and the importance of subjective aspects of emotion are real issues, it is possible to defend the methodological choices made for these studies. Individual differences in SCR and SCL are indeed large, but the transformation procedures used in these studies correct for them, even if they are incapable of forcing physiological data to fit parametric assumptions. It is not exactly fair to argue that large individual differences necessarily lead to inaccuracy – people vary greatly in their musical taste, this hardly means the notion of musical taste is irrelevant. The relevance of physiological indices of affect may be contested by affect theorists, but it should be noted that chapters 7 and 8 included incidental self-report measures as well as physiological measures and both measures yielded nearly identical results.

It is also worth noting that only *trait* EI tests were used in this thesis. The null correlations between EI test scores and reactivity to affect induction (i.e., affect regulation) may be due solely to the properties of the EI test. If the test does not measure affect regulation particularly well, it is perhaps unsurprising that results on it do not correlate with other measures of affect regulation. The primary reason why *ability* EI tests were not used here is because of the expense associated with using such measures and the inability to examine raw item data. The study carried out by Ciarrochi et al. (2000) used *ability* measures, but it would still be of use if future research could examine whether or not there is a relationship between *ability* EI test scores and reactivity to affect induction.

These findings are limited in scope. As was introduced in chapter 6, only one physiological indicator of affect was used in this thesis: skin conductance. There are numerous other physiological indicators which could have been used and it is entirely possible that EI tests predict activity in other physiological correlates of affect (e.g., heart rate, pupil dilation, etc) even if they fail at predicting SCR or SCL.

It would be irresponsible to over-generalise these findings and claim that EI theory does not work simply because the indices used in this study do not correlate with EI test scores. The most sensible conclusion seems to be that the results of these experiments do not support current EI theory but it is possible that these results are simply due to a limited (i.e., only SCR, SCL, and self-report indices of affect were used) number of indicators of affect. The use of a different outcome measure in chapter 9 partially addresses this limited scope, but it is still accurate to say that the findings in this thesis only investigate a limited theoretical area.

10.3 Relevance of Findings to Current EI Theory

The findings of these studies do not fit well into what would be expected according to either *trait* or *ability* EI, though the difficulty is more pronounced for *ability* EI. The psychometric and experimental studies seem to indicate that current EI theory does not make predictions that are borne out in the data. These findings are immediately relevant to this theory but also lend themselves to more fine-grained analysis and it is to this aim that the discussion will now turn.

The results of chapter 4 can be interpreted in a couple of ways in light of current EI theory. As was noted above, the method of interpreting the results favoured here is to accept the findings of the tests at face value and to argue that the results indicate that responses to EI tests are more similar to what would be predicted according to the conjecture than according to current EI theory. This interpretation was favoured over the more conservative interpretation that merely states that EI tests are flawed because discarding EI tests out of hand dramatically limits the possible areas of discussion. It is not possible to fully endorse both interpretations simultaneously and although it would be tempting to argue that the EI tests are inaccurate for measuring current EI and not 'conjecture EI', this argument is neither fair for consistent. Both of these interpretations call current EI theory into question to different degrees: either EI tests are flawed and thus so is the theory which prompted them or the tests are acceptable and the results from them supports an alternative theory. Either way, the findings suggest current EI theory leaves something to be desired.

Yet it has been consistently reported (e.g., Austin, Saklofske, Huang, & McKenney, 2004; Petrides & Furnham, 2000) that EI tests have unclear factor structures which vary between reports, so it may be claimed that differences in factor structure do not contraindicate EI theory as much as it might initially seem. Although the findings in chapter 4 were fairly robust, it could be that EI tests simply need a great deal of revision and any issues they have should not be taken too seriously. Unfortunately, this defence runs headlong directly into the argument that EI tests are flawed and thus so is the theory that prompted them. That is, if EI tests are so flawed, then it is unclear why they are used and why *any* findings (whether in the predicted direction or not) should be taken seriously. In other words, an attempt to defend EI theory from the findings in chapter 4 by pointing out that EI tests have difficult factor structures fails because of the second interpretation. It should be noted that the EI tests used in this study are fairly new so the findings of chapter 4 should probably be interpreted generously. It is not possible to entirely dismiss the 'bad' findings from chapter 4 (unless one also wanted to dismiss the 'good' findings as well) on the grounds that EI tests are new, but it is probably appropriate to interpret the results of that chapter leniently.

It may also be the case that the differences in factor structure found in chapter 4 could be incorporated into current EI theory. It should be noted that these measures are *trait* measures designed to asses a collection of self-beliefs and behavioural tendencies and not necessarily skills. It is not entirely necessary that a measure of self-beliefs have identical factor structures in high- and low-scoring groups because it may be that these respective groups' self-beliefs or behavioural tendencies are meaningfully different and factorial differences are merely detecting a real difference. Of course, if EI is viewed as an *ability,* more difficulties present themselves. It would need to be made clear which mechanisms or properties of EI skills could cause them to be so different from normal skills and abilities. A normal skill (e.g., penalty kick ability) can be measured on one dimension (e.g., % of attempts made) and it is unclear what could cause EI skills to require two or more dimensions. In fact, it would seem that any attempt to explain these findings in the frame of *ability* EI theory might lead to something similar to the conjecture rather than a straight-forward integration into current EI theory.

The results of the remaining chapters also seem to present a difficulty to current EI theory. In all but one comparison it was found that EI test scores do not correlate with indices of affect regulation when the entire samples were analysed and when participants were given clear, explicit instructions to regulate. Yet it is possible that part of individual differences in affect regulation ability is an individual difference in how 'necessary' it seems to regulate. The stimuli were selected based on how they tend to make participants feel (IAPS ratings of valence and arousal) but perhaps inducing a particular affective state is an entirely different matter from prompting affect regulation. Perhaps people with good regulation abilities differ from one another in how necessary they feel regulation is. They might all be affected equally but differ from one another in whether they feel that it is 'worth it' to regulate. Similarly, it is possible that multiple EI skills compete with one another for the dominant effect. As previously suggested, Petrides and Furnham (2003) highlight a 'sensitivity' effect of EI and Ciarrochi, Chan, and Caputi (2000) highlight a 'control' effect, both of which lead to opposing findings. Perhaps something similar is at work in the studies of this thesis. Perhaps high-EI individuals are both more 'sensitive' and capable of 'controlling' their affect, but these two abilities compete against one another in a way that obscures any overall correlation between EI test score and evidence for affect regulation. Either of these suggestions would complicate EI theory somewhat but might be able to make sense of these findings.

The possibility that 'control' and 'sensitivity' compete and the responses of participants defy prediction due to this competition seems unlikely given the constraints of these experiments, but it may still be the case. Current EI theory posits that both sensitivity and regulatory ability are part of EI and either or both could be in play in any given situation. Perhaps the null findings of these studies are not due to any failure of current EI theory, but rather reflect the *correct* functioning of two conflicting skills simultaneously. If each stimulus prompted each participant to make a choice between two competing responses then a great deal of noise in the data is hardly surprising. This state of affairs seems unlikely, at least in chapters 6-8 which included a 'clear' condition in which all participants were requested to neutralise their affect. Surely if EI is comprised of both the ability to neutralise and the ability to be sensitive, then an explicit request to neutralise should override the sensitivity reaction and thus a neutralisation effect should be observed. This was not the case and as such it seems unlikely that this 'opposing skills' proposition is likely.

Regardless of whether or not any particular explanation is successful, it is also worth discussing the prospect of theoretical re-explanation more generally. The findings of this thesis stand at odds with what would be predicted based on current EI theory and while it may be the case that theory could be expanded or altered to suit the data, such theoretical adjustments may be considered dubious by some. Surely any expansion of current theory must be congruent with the existing findings – it would not be acceptable to add something such as murderous tendencies to any theory of EI – so the suggestions above and in general any theoretical expansion must make sense in light of current EI theory. These 'competing' or 'hierarchical' skills may well make sense, but it would be necessary to show that they make sense in light of the rest of EI theory. More importantly, any *post-hoc* explanations must also make testable predictions. These requirements may seem pedantic or obvious but fairly grave mistakes have been made in the past when attempts to justify theories in the face of contradictory evidence.

It also seems worth nothing that altering current EI theory by adding additional layers would be similar, at least in spirit, to the conjecture. As was noted in chapter 3, the conjecture is itself an extension to EI theory rather than a replacement. The suggestions for theoretical expansions above are only two amongst any number of possibilities, including the conjecture. Certainly these various suggestions differ from one another, but they are all similar in the respect that they may account for why EI theory fails to make accurate predictions in this thesis. If a theorist is open to the enterprise of suitably altering EI theory in one manner, then it seems only reasonable to at least consider the possibility of others. In other words, if the possibilities above are entertained, then the conjecture is probably also worth considering, especially as it makes some limited number of accurate experimental predictions.

Chapter 9 did not asses any hypotheses based on current EI theory specifically because it was already certain that the data analysed in that study would support current EI theory. However, this is not any attempt at duplicity. Hypotheses based on current EI theory were not tested in this study simply because the data examined in chapter 9 had already been presented as support for current EI theory by Saklofske et al (2007). This indirect support seems to favour existing EI theory and is discussed below, this chapter provided no support for the conjecture.

The results of these studies provide similar amounts of support for current EI theory as previous studies have. Previous experimental studies provide limited support for EI. Ciarrochi et al (2000) found that only some of their predictions were borne out and while Petrides and Furnham's (2003) findings supported their predictions more robustly, Austin (2004) and Farrelley and Austin (2005) found limited support for EI theory in that approximately a quarter of the positive predictions were supported. The studies in chapters 4-8 found only one significant correlation between EI test scores and indices of affect regulation out of more than ten comparisons and thus there is little evidence in favour of current EI theory. It simply *must* be suggested that perhaps EI theory needs be seriously adapted based on the findings of this thesis. It is difficult to know what to do with a theory which makes accurate psychometric predictions (see introduction) with tests that do not exactly test what they are designed to (chapter 4) but which makes few accurate lab-based predictions. A strict, conservative researcher might argue that EI theory should be discarded because of the poor experimental evidence in its favour. This may be overly harsh as there does seem to be *some* evidence that EI generates accurate experimental predictions, but it is not transparent how much evidence is necessary to decide that EI is an experimentally valid construct. Of course, there are no hard and fast 'rules' in science regarding what constitutes support for a theory. However, it can probably be safely claimed that this thesis, which found support for only one hypothesis out of over ten, provides very limited evidence in favour of current EI theory. Perhaps Austin's (2004) finding that only a specific subscale of EI tests predicted emotional inspection time performance can be constituted as support even if this significant finding came along with several nulls, but the studies here do not even provide that much support.

To conclude, the studies of this thesis have immediate relevance to EI theory. The obvious interpretation is that these findings cast some doubt on current EI theory as very few of the predictions that were made were actually supported. It may also be possible that if EI theory is suitably expanded or slightly altered then it could account for these findings, but the conjecture is probably not suitable to explain these findings given that the conjecture received so little support. It should also be repeated that these findings are necessarily limited in scope to physiological aspects of arousal only.

10.4 Relevance of these Findings to the Conjecture

The *awareness* and *flexibility* conjecture is the suggestion that part of EI is heightened *awareness* of the variety of ways an individual emotional situation could be resolved and greater *flexibility* in enacting these various solutions. The results of the studies in this thesis provide very limited support for this conjecture in that a fair number of the hypotheses which it prompted were supported. This support was limited at best and there are numerous theoretical issues worth discussing.

The results presented in chapter 4 support the hypotheses driven by the conjecture rather than the hypotheses generated according to current EI theory but these findings could be construed in other ways. It was found that the dimensional structure of EI tests differs between EI subgroups and this was what was predicted based on the conjecture – high-EI individuals were expected to answer EI questions in a different manner (i.e., above and beyond simply scoring higher) than low-EI individuals. Specifically, it was found that factor loadings in different sub-groups were only weakly correlated with one another and this finding, effectively a null, was contrary to the strong correlation expected based on current EI theory and classical test theory. However, as was discussed above, using a null finding to support a theory is difficult. The lack of correlation between factor loadings could have been due to any number of factors and the nature of null findings makes it difficult if not impossible to determine what exactly the cause was. As such it could easily be argued that the null findings do not necessarily support the conjecture and there is no responsible defence from this criticism – such is the risk that is run when a null is predicted. However, the findings of this study do illustrate how the conjecture might have been able to explain a finding that does not fit well into existing EI theory.

Chapter 5 was the first study to make any positive (i.e., non-null) predictions based on the conjecture and the hypotheses were both supported. The conjecture leads to the expectation that people low in EI (i.e., below the mean) would regulate their affect in line with their EI level whilst people high in EI would regulate their affect unpredictably. This simple test of the conjecture did indeed show the expected results. According to scatterplots of the data, it appeared that for sub-mean scorers, increased EI had a neutralisation effect. That is, participants with very low EI scores were greatly affected by the affect induction (and in the direction expected – i.e., the negative film caused an SCL increase), middling scorers had virtually no change in SCL at all, and high-EI individuals' SCL changes were erratic and scattered. In other words, for sub-mean scorers, increased EI was associated with smaller changes (i.e., neutralised) in affect but this relationship disappeared for the high-EI scorers. This is exactly what was predicted according to the conjecture. This finding seems to free the conjecture from the trap which it had fallen into in chapter 4: it is no longer necessary to draw support from nulls alone. The positive predictions made by the conjecture in this study were supported by the data as were the nulls and although it is possible that it was due to a lack of any explicit request to neutralise affect, this possibility is explored in later chapters. It would seem that although this experiment was somewhat unsophisticated, it did support the conjecture.

Chapters 6 through 8 introduced more complexity to the studies. As was explained in chapters 4 and 6, ambiguity is central to the conjecture because the *flexibility* aspect of the conjecture is only applicable when multiple solutions are present. Choice in response is essential for *flexibility.* However, it was also necessary to include a 'clear' condition which explicitly requested that participants neutralise their affect for reasons also described in chapters 4 and 6. The inclusion of these two conditions provided a way to test current EI theory and the conjecture simultaneously without necessarily placing them in opposition to each other. Chapters 7 and 8 were replications of chapter 6 which also included incidental self-report measures of affective state to serve as a manipulation check and as a tentative exploration of whether the hypothesized relationship between EI test scores and reactivity to affect (i.e., affect regulation) existed in self-report data.

The results of chapter 6 only allowed the rejection of the null hypothesis in favour of one positive hypothesis based on the conjecture – the predicted relationship between EI test scores and affect reactivity occurred in the low group for the negative image ambiguous condition. The predicted nulls were found as well, but it is probably more sensible to focus on the positive findings. It is unclear why this neutralisation effect only occurred for the negative images. The conjecture suggests that low-EI individuals are characterised by predictability in their responses to stimuli, but perhaps even low-EI individuals are aware that positive stimuli and positive feelings could prompt numerous reactions. It isn't hard to imagine any participant thinking 'I feel happy, why should I change this?' Or perhaps the combination of a lack of any explicit instruction *and* the increased ambiguity of positive stimuli interacted to prompt even low-EI individuals to behave in a variety of ways. It may also be that the IAPS images used to induce positive affect functioned so variably between individuals (e.g., an image of waterskiing might have an opposing or negligible effect on two different people) that *flexibility* was ensured simply by the individual variation in reaction to the stimuli, let alone variation in neutralisation. Regardless of the reasons behind the failure of one experimental hypothesis, the correlation between EI scores and SCR in the negative images was as expected. More generally, the main experimental manipulation between 'clear' and 'ambiguous' seemed very important as the correlations that were present in the ambiguous condition disappeared entirely in the clear condition. Thus the findings of chapter 6 seem to support the conjecture in an immediate, clear manner, albeit more weakly than the findings of chapter 5.

Chapters 7 and 8 yielded very similar results but the incidental tests of the self-report data resulted in few cases in which the null hypothesis could be rejected. The studies in these two chapters were nearly identical to chapter 6 save that self-report ratings of affective state were collected along with SCR data and participants. The findings of these studies varied from one another in some respects but on the whole both yielded results which allowed the null hypothesis to be rejected in favour of the conjecture-driven hypotheses in some cases. In both studies a greater number of correlations were predicted than were found, especially in the self-report data which mirrored the findings of chapter 6.

Chapter 9 used a different method for creating the high- and low-EI subgroups, used a different dependant variable, and used an additional operational definition of *flexibility.* This study of the relationship between EI test scores and coping styles found no support for the conjecture-driven hypotheses and this seems to indicate that whatever limited support found for the conjecture was due largely to methodological decisions and not a real effect.

It is difficult to know precisely what to do with findings that support only some of the predictions made and what remains is the decision between two less than ideal alternatives. It is hard to conceive of a reason for why the neutralisation effect would occur in different experimental conditions in different studies and the fact that the effect was found in the positive *and* negative conditions in different studies seems to rule out the explanation based on the notion that positive stimuli are too ambiguous. Even if these unpredictable findings can be explained it is still difficult to ascertain how much support these studies provide for the conjecture. Certainly *some* of the hypotheses were supported (roughly half in the physiological data) but it is not clear whether or not this is sufficient evidence to accept the conjecture. It is also uncertain whether a halfway supported conjecture would make for a sensible addition to existing theory.

As was explained above, it is normal for a study to be arranged such that hypotheses based on two theories are mutually exclusive, but even though this was not done in studies 6 through 8 it is still possible to use the standard 'selection based on alternatives' method to arrive at a conclusion about the conjecture. It should be noted first of all that in chapters 4 and 5 the conjecture-driven hypotheses were supported at the cost of those hypotheses based on current theory. Certainly, the fact that in chapters 6 through 8 the hypotheses based on the conjecture and current EI theory were not mutually exclusive makes a decision slightly less clear – it is possible that both theories could have been unsupported or supported – but because of the way the results *did* emerge, it seems clear that current EI theory was not supported and the conjecture received very limited support. Of course, the goal of the conjecture was not to replace current EI theory but to be accepted as a useful theoretical addition, but even this limited aim did not receive enough experimental justification.

If nothing else this thesis highlights the importance of ambiguity in EI research questions, though this additional factor should not be too surprising. It may be worthwhile to further examine the finding that ambiguity is a critical undercurrent in EI research. Chapter 4 showed that EI tests which use items that represent a multiplicity of situations (i.e., are ambiguous) are psychometrically and theoretically problematic for current EI theory. Chapters 5 through 8 showed that in an ambiguous situation there are important differences in how low- and high-EI individuals use their EI skills and that including an explicit, clear task, reduced these differences. There doesn't appear to be any way to avoid the fact that ambiguity and the use of EI-related skills are intimately linked. This should not surprise anyone given the self-evident ambiguity of emotional situations and problems. As was explained in chapter 3, most emotional problems have numerous solutions and as has been noted since the birth of EI theory (see Matthews, Roberts, & Zeidner, 2002), the 'no correct answer' problem is a major stumbling block for measuring EI. The results of these studies seem to suggest that the 'no correct answer' dilemma is also important in the lab. This difficulty is hardly insurmountable and can be incorporated into either current EI theory or the conjecture fairly easily. The relevance of ambiguity to the latter has already been explained several times and it seems to slot into current theory simply by noting that it is hardly surprising that participants behave unpredictably in the absence of a clear task. The entire purpose of instructions is to ensure that participants engage in a certain behaviour to the best of their ability and it is hardly contrary to current EI theory to show that participants behave unpredictably when they lack instructions, especially in something as novel and ambiguous as an SCR affect induction paradigm. As unsurprising as the importance of ambiguity is, it is still noteworthy. Future studies of EI-related skills will need to be carefully designed with this finding in mind.

It should also be clarified again that the conjecture was not introduced to refute current EI theory, merely to expand upon it. There are certain sub-skills which seem to not fall within the purview of the conjecture. That is, it is hard to conceive of a way in which a skill such as emotion recognition would be subject to *awareness* or *flexibility,* rather, high-EI probably just results in greater ability to recognise emotions.

10.5 Conclusion

The most responsible conclusion seems to be that the conjecture was not well supported enough in these studies to be accepted as a useful theoretical addition to existing EI theory. The findings of this thesis also fail to provide any notable amount of support for existing EI theory.

Chapter 11: General Discussion 2 – Experimental Limitations, Recommendations for Future Studies, and Conclusion

11.1 Limitations

There were flaws in several of the studies in this thesis, some of which were addressed by subsequent studies and some of which were not. Some were legitimate methodological choices which can be defended and some are flaws which should be corrected in future work.

11.1.1 Chapter 4

One of the flaws of chapter 4 was that a limited assortment of EI tests was used. Although all of the tests were multidimensional, widely-used EI tests, there was only one *ability* EI test used and most of the measures were short measures (i.e., 50 items or less). In a way, our 'sample' of EI tests only represented fairly short EI tests. It is possible that the findings of this chapter do not apply to longer tests of EI. This seems somewhat unlikely because the tests used in that chapter represent three different formulations of EI theory (Salovey & Mayer's (1989) original three-branch theory, Mayer, Salovey, Caruso, & Sitarenios' (1999) four-branch theory, and Bar-On's (2002) five-branch model) with four different tests. Simply put, the findings of chapter 4 were robust across nationalities and tests, so it seems likely that we can generalise these findings to other EI tests.

Another issue with this study was the lack of any deeper investigation into the causes of the factorial variance. It was argued that the dimensional variability between EI subgroups was due to the increased action of *awareness* and *flexibility*, but there was no way to be certain from these data. Aside from the complex logic presented in that study, there is no clear, let alone convergent, indicator that participants were indeed expressing more *awareness* or *flexibility*. Of course this study was designed to be an introduction rather than the culmination of several studies. Psychometric findings are often considered less convincing than experimental findings.

Additional minor flaws are discussed in that chapter, but it is probably accurate to say that this study is more informative about the psychometric properties of EI tests than it is about EI theory or the conjecture. This study also highlights the importance of experimental studies in evaluating psychometric claims.

11.1.2 Chapter 5

The flaws of chapter 5 are discussed in that chapter but there are issues of more general relevance as well. The film stimuli used in this study could have resulted in some kind of demand characteristic. It was certainly clear from even the first several seconds what kind of affective reaction the film clip was selected to elicit and it is possible that participants only reacted to it in the expected manner because of this cue. It seems unlikely that such a demand characteristic could happen. Although it is possible to 'fake' skin conductance readings (by clenching muscles, breathing rapidly, etc), doing so requires conscious effort and by definition, demand characteristics are unconscious, so for a demand characteristic to be problematic, it would have to somehow convince the participant to actively 'fake' his responses. Even if this were possible, the question of motivation presents itself. It is not clear why a participant would actively 'fake' his/her responses, whether there was a cue present or not, given that they were not being rewarded for the 'correct' response, nor were there any tasks to be completed during or after the stimuli. Even if all of these impediments to demand characteristics (i.e., difficulty in 'faking', the crossing from unconscious cue to conscious 'faking', and motivation) were overcome, it must be recalled that this was a study of affect regulation. Even if participants were being subtly directed to emotionally react in the desired manner, the study examined changes from baseline over the course of a minute – i.e., relative rather than absolute values. That is, for such a cue to be problematic, it would have to cause participants to somehow react such that their changes from baseline were systematically different in different film categories. This seems highly unlikely.

More globally, the method of affect induction used in this study could be called into question. A strident defender of current EI theory might suggest that 'true' EI operates only in less constrained situations. That is, the argument might go that because this study used such a limited type of affect induction (film stimuli) it failed to accurately represent actual affective situations. Or perhaps in order for EI to function properly, the stimulus that induces affect must include more social content or context. In other words, this study could be criticised for lacking sufficient ecological validity to accurately reproduce the conditions which are necessary for EI skills to be observed. This complaint seems less than compelling. EI is often claimed (e.g., Bar-On, 2002) to be a robust, global set of skills which make a powerful impact in quality of life and in observable behaviour. It seems contradictory to claim that EI skills, self-beliefs, or behavioural tendencies (e.g., affect regulation) are robust and global yet these same skills, self-beliefs, and tendencies can only be observed under certain experimental conditions. Surely if EI is robust, it should be possible to observe its effects under a variety of experimental circumstances. Additionally, it is a truism of psychological research that ecological validity stands in opposition to experimental precision. As was explained in chapter 1, the aim of this thesis was to use a precise, 'objective' measure of affect in a strict experimental paradigm. So even if this study lacks ecological validity, it is considerably more precise than would be possible in something like a field study or an experiment in which affect was induced by a confederate or something similar. Certainly other means of affect induction should be used to test EI theory or the conjecture, but the fact that this particular technique resulted in poor support for current theory can not be mitigated by an appeal to ecological validity. It is possible that EI skills only operate under certain affect-inducing conditions, but there is no real *a priori* reason to believe this is true given that is usually claimed to affect all areas of life (otherwise less stringent methods would have been used in these studies) and such an argument seems more of a distraction than a real criticism.

Another major issue with this study as well as chapters 6 through 8 is the arbitrary nature of the high/low group split and the importance this split makes to the hypothesis test. If the groups are divided had been divided by a cut-off score which was only slightly different, the pattern of results would have been drastically different. This arbitrariness seems problematic but it does not seem catastrophic. The purpose of these studies was a preliminary test of whether or not the conjecture had any currency. Put simply, the groups had to be cut *somewhere* and a point on or near the mean seemed sensible. Median and mean splits are very common when psychometric tests are used to divide participants into sub-groups. The purpose of these studies was to set up experimental conditions which would allow the conjecture to function if it has any utility whatsoever. Groups were split according to a score near the mean to divide the sample roughly in half but some flexibility was employed in the group splits to provide conditions generous enough for the conjecture to succeed. Obviously, the failure of the conjecture to make accurate predictions even under such generous conditions is informative.

However, the median/mean split approach is less than ideal because the group divisions are entirely sample bound and there is a lack of any external justification of these division criteria. An extreme-groups approach would have been preferable because it would have allowed some certainty that the subgroups' EI scores were meaningfully different from one another. This is the approach used in chapter 9.

11.1.3 Chapters 6 through 8

One issue with these studies was that the hypotheses driven by the conjecture and current EI theory were not written to be mutually exclusive. Rather, the ambiguous condition was included to test the hypotheses based on the conjecture whilst the clear condition was designed to test the hypotheses based on current EI theory. The result of this arrangement was that it was possible for both groups of hypotheses to be supported (or unsupported) at the same time. That is, it would be possible with this arrangement to have *both* theories supported or not supported at the same time. Were this outcome to be observed, it would not be possible to rule out one theory in favour of the other through the normal 'selection by rejection' process. As was explained in the previous chapter, it is still possible to arrive at conclusions based on selecting from two alternatives even without a set of hypotheses arranged to be mutually exclusive, but it is also worth explaining the arrangement which was used.

The hypotheses were arranged in such a manner because it was not the purpose of this thesis to replace existing EI theory with the conjecture. The arrangement used in chapters 6 through 8 was essentially two experiments combined in one – the ambiguous condition tested the conjecture and the clear condition tested current EI theory. Each condition made positive predictions according to a theory and tested them, so in effect each condition was an independent test of a theory. By arranging the study in this way it was possible to test both theories at the same time and by comparing them to their own respective nulls rather than to each other directly. This seems appropriate given the incremental (rather than 'confrontational') nature of the conjecture. Perhaps this arrangement was less than ideal, but given that the purpose of the studies was not to replace EI theory, but to augment it, it seems appropriate. At no point was it argued that the conjecture would replace existing EI theory.

It should also be noted that even though the experimental logic was not arranged such that the two theories were mutually exclusive to one another, it is still possible to decide between them. The two theories were assessed by rejecting the nulls which accompanied each experimental hypothesis and this seems like it should be sufficient. It seems reasonable to compare two theories based on whether or not the null hypothesis could be rejected in favour of a positive prediction made by that theory. It might also be acceptable to compare theories based on the ratio of times the null hypothesis was rejected to the number of overall predictions made. This manner of assessing the theories allows each to be measured according to its own merits rather than necessitating a choice between them.

Another issue with chapters 6 through 8 was the specificity of the operational definition. As was discussed in chapter 1 and above, emotion is a vast, complex behaviour and the use of such a strict definition of affect in this study is a limitation.

Although chapters 7 and 8 also included subjective ratings of affective state as an incidental check of whether or not the EI-reactivity link existed in subjective data as well, the chief interest of this thesis was physiological indices of affect. Perhaps participants' affective reactions were being misrepresented or over-simplified by the SCR measure. SCR readings were only taken for a five-second duration and as such it is possible that the 'window' of observation missed some important aspect of the physiological response. It is unlikely that the five-second window missed important findings as even this limited window was sufficient to see participants' reactions to the stimuli as well as their gradual habituation to the setting. If observations had been made for ten seconds, very little additional information would have been gained and it is likely that a more extreme relaxation trend would have been observed. More to the point, as was mentioned above and in chapter 1, this thesis was specifically interested in if and how EI skills are observable in a precise laboratory setting. Previous studies (e.g., Ciarrochi, Chan, & Caputi, 2000; Petrides & Furnham, 2003) used subjective methods of affect induction and affect measurement, but this thesis was designed to address the dearth of concrete, 'objective' evidence of EI. It would be difficult to thoroughly investigate EI using a precise laboratory setting *and* a more ecologically valid or observational study in a single set of experiments. It is also hardly a flaw that this study used a well-validated, highly precise measure of affect. It is certainly possible that the measure used here missed out on some facets of the affective response, but this would be the case with any single measure of affect. Unless every possible index of affect was recorded, the 'you missed a spot' argument could be made, so although it is a valid complaint, there were legitimate reasons behind the use of physiological measures in this thesis and because this may be the first SCL/SCR study of EI, it seems unduly harsh to apply this criticism.

11.1.4 The Conjecture and Certainty

One of the major flaws of all of the studies is that there was no independent check of *flexibility* or *awareness*. This was noted briefly above but is worthy of deeper discussion. To put it simply, these studies have been designed such that variability in responses has been construed as flexibility but there is no guarantee that this should be the case. This is a similar issue to the 'support from null findings' issue discussed in the previous chapter.

In chapter 4, dimensional variability was taken as evidence for the increased action of *flexibility* and *awareness* and while this is certainly a reasonable interpretation it is not the only one. It is possible that some other theory could explain why the EI test responses of different EI subgroups had different factor structures, despite the obvious expectation that the same factor structure would run through all responses. In fact, some other explanations were presented in the previous chapter. These possibilities or others are certainly interesting and worthy of investigation.

The issue with these experiments is not that the operational definition of *flexibility* was behavioural or psychometric variability. This is a perfectly defensible definition given that the very word 'flexibility' entails by definition variability. Just as lower reaction time is used to operationally define processing speed, greater behavioural variability can be used to operationally define *flexibility*. Moreover, chapter 9 provides an additional operational definition of *flexibility* as increased use of highly situation-dependant skills when more generally useful skills have already been employed. The issue in these studies is not the way a single operational definition was used, the flaw is slightly more subtle.

What would have been ideal is a second definition of variability in each study because the operational definition used is self-referential. Any study which uses a single operational definition is in some way tautological and the way most researchers avoid this problem is by running many studies each of which use slightly different operational definitions. Although chapter 9 and the inclusion of self-report measures in chapters 6 through 8 address this flaw, this thesis would have been stronger for having a greater number of operational definitions. Also, the operational definitions in these studies are functional, but they are not particularly well-developed. It is not clear how either of these aspects of the conjecture functions, how it may differ between participants, or how it relates to overall EI, amongst other issues. This is a problem which should be addressed in future studies. Although the theory behind the conjecture is explained in chapter 3, the explanation is purely in the abstract and does not address the possibility that the conjecture was not responsible for the findings in these studies. It is certain that if some sort of provisional support was granted to the conjecture, future experimental studies would be necessary to determine if there are better ways of operationalising the conjecture. It can only be said that the operational definitions here do indeed function across numerous different studies, so although it is certainly possible that something other than the conjecture could have caused these results, they are also congruent with what the conjecture predicted.

Another issue regarding certainty was the potential for type I errors in these studies. Especially in later chapters, multiple correlations coefficients were calculated and it is possible that the few that were found were simply due to chance. A simple Bonferroni correction would reduce all of the findings here to non-significance, but alternative corrections would be provide a balance of rigour and power. Of course, it has been concluded that the conjecture is probably not worthy of future investigation, so the possibility of type I error merely reinforces this conclusion.

11.1.5 Sampling

Another limitation to this thesis was the use of students for samples. This is a problem that a large number of psychological studies faces and it would be tempting to simply wave away this criticism for the fact that it impacts most studies equally, but this issue is of particular import for EI studies.

Many times the 'sophomore problem', the issue of using university undergraduates (many in their second year) can be ignored because it is difficult to conceive of a way in which undergraduates differ from anyone else, but in the case of EI, this problem may not be so easily ignored. Because EI is such a socially relevant set of skills and because it is posited that EI increases with age, it is possible that EI differs meaningfully between undergraduates of different ages, degrees, or social activity. Because the undergraduate experience is a tumultuous mix of socialisation, personal exploration, and personality development, it is plausible that undergraduates are meaningfully different from people who are already settled in careers. Thus it may not be accurate to infer that findings based on undergraduate behaviour are representative of how non-undergraduates behave. It is often assumed that findings from studies that use undergraduates can be extrapolated to the normal population but this may not be the case for this thesis.

However, it is also possible that undergraduates do not meaningfully differ from other young people. There is certainly something unique about the undergraduate experience in that it involves education, but in all likelihood, most if not all people aged eighteen to twenty one go through a similar period of personal exploration, socialisation, and personality development – even if this period lacks lectures. Simply put, it seems unlikely that undergraduates differ from non-undergraduates of the same age in terms of their EI. So perhaps the real issue is age-related, as it has been reported that EI test scores are positively associated with age (Van Rooy, Alonso, & Viswesvaran, 2005). It seems more likely that there is a meaningful difference in EI between old and young people. An issue related to this will be mentioned below in the recommendations for future research.

It could also be claimed that the sample sizes in this thesis were unacceptably small, at least in some chapters. Specifically, the sample size in chapter 7 was very small due to the equipment issues and outliers that were only noticed after the study was completed. Chapters 5,6, and 8 had larger samples and although it is always possible to demand larger sample sizes, these seem to be reasonable. Large samples would have increased statistical power and this would have had important ramifications for these studies. The effect sizes here are doubtless smaller than those found by Petrides and Furnham (2003) who also had a small sample ($N = 30$), but the marginally significant effect sizes in these studies (i.e., $\rho = .3\text{-}.5$) seem large enough to be significant in a larger sample. Larger samples might also have helped correct for the variability in physiological responses between participants. That said, the findings were fairly straightforward even if some studies' results contradicted others so it is unlikely that a larger sample would have drastically altered the findings. Of course, it is always desirable to have more participants take part in a study until the number reaches a point where it virtually guarantees significance.

11.2 Recommendations for Future Studies

11.2.1 Psychometrics

There are numerous avenues for psychometric investigation of the conjecture which expand upon the results of chapter 4. From a raw psychometric perspective, it would be worthwhile writing a measure which could be used to assess the conjecture. Responses to this measure could be correlated with criteria to determine if it has any predictive validity. The psychometric properties of any test based on the conjecture would be of great interest. It would be interesting to find out whether or not tests based on the conjecture could indeed pin down such a seemingly slippery concept as *awareness* or *flexibility,* what kind of factor structure such a test had. The scales developed to measure self-monitoring may be a useful model for testing *flexibility* or *awareness,* as would the measure designed by Cheng (2001) to measure coping flexibility. Another worthy avenue would be to replicate chapter 4 using more sophisticated statistical procedures capable of determining if dimensional variability occurs after intelligence and personality variance has been controlled for. Another obvious question is whether or not normal EI and conjecture-based EI correlate at all as it would seem unexpected if they did not. In other words, future studies could examine whether or not the conjecture passes all of the tests that we expect every psychometric construct to pass.

11.2.2 Affect Induction

There is even more room for future studies of how affective reactivity relates to EI. The physiological paradigm used here could be retained but a confederate could be used to induce happiness or distress in the participants. Or the same physiological recording could be taken while participants were presented with a longer, more complex stimuli (such as a conversation with a confederate or a film of a conversation) with precise time-stamped events which could be cross-referenced with specific SCRs. A myriad of possibilities present themselves as nearly all previous research in EI has been psychometric rather than experimental.

The lack of prior research makes it difficult to recommend one affect induction method over another but it may be the case that the affect-relevant images used in these studies are inappropriate. Certainly the results from this thesis seem to suggest that using discrete images might result in little support for current EI theory but perhaps this is due to the images rather than the theory. This seems somewhat unlikely as the images used all had previous ratings of affective impact, but as was mentioned above, perhaps an image simply does not contain the necessary context and content to successfully induce affect. Or perhaps unlike film stimuli, images do not contain enough of the cues necessary to induce affect properly. Or perhaps the images were successful in inducing general anxiety or happiness but these rather vague feelings were not sufficiently specific to be recognised as actual affective states (e.g., happiness) by the participants. Any of these possibilities could be used to explain why the results of chapters 6 through 8 resulted in such poor support for current EI theory even if they failed to explain the results in chapter 5. Regardless, it is probably worth investigating what role induction method plays in outcomes. A study which contained two types of affect induction in the same paradigm would be useful in determining whether or not the induction method affected the results.

11.2.3 Other Experimental Arrangements

Owing to the fact that there have been so few non-psychometric studies of EI, there is a great deal of room for future studies of affect regulation. It would be entirely feasible for an experimenter to have a confederate deliver some sort of distressing or elating news to participants of different EI levels and for observers to rate the participants' reactions. Or in an even less lab-bound study, an observation of how high- and low-EI individuals behave in arguments or in jocular situations would be revealing. A study which used subjective indices of changes in affective state (i.e., control for baseline affective state in a manner similar to SCR or SCL) but was otherwise similar to chapters 6 through 8 could be worthwhile as it would augment the incidental indices used in those chapters.

There are also other EI-related skills which could be investigated using experimental methods. At present, experiments have been largely concerned with affect regulation (e.g., Petrides & Furnham, 2003) or facial emotion display recognition (e.g., Austin, 2004). There are at least two other EI-related skills: emotion utilisation and emotion comprehension. Under Bar-On's trait theory there may be up to thirteen more component skills, self-beliefs, or behavioural tendencies. It should be straightforward to measure some of them (such as the 'positive self-regard' facet of Bar-On's theory), but it would initially seem that some of these skills (e.g., emotion utilisation) would be difficult to operationally define for experimental study. Yet it would certainly be possible to fit these seemingly vague skills in the laboratory – emotion comprehension ability could be measured as the number of 'meanings' associated with emotions, for example – and it would be important to do so. At least half of the skills associated with EI have not been examined in the lab and as was pointed out in chapter 1, experimental evidence is critical if a construct is to be taken seriously.

11.2.4 Other Types of Flexibility and Awareness

As was briefly noted above, this thesis used a somewhat limited number of operational definitions for *awareness* and *flexibility*. Although circular logic was avoided by making different predictions for high- and low-EI individuals there is still a lack of variety in the sorts of behaviours which could be taken as evidence of *awareness* and *flexibility*. One potentially interesting avenue would be interview data. That is, given that this conjecture is as yet only an interesting possibility, it would be interesting to see how high- and low-EI individuals describe their approaches to emotional situations. It is possible that high-EI individuals would provide more numerous and more sophisticated responses to situations. It might even be fruitful if such interviews were combined with the aforementioned attempt to create a measure of *awareness* and *flexibility*. A similar, ecologically valid operational definition for the conjecture would be negotiation skill. A study could be designed where high- and low-EI individuals (either selected with current tests or a hypothetical conjecture EI test) were asked to negotiate for a desired outcome or object either with a confederate or another participant. Certainly there are numerous other ways in which the component parts of the conjecture could be measured. These avenues were not explored in this thesis because it seemed a wiser course of action to look at EI and physiological reactions in depth rather than to examine several operational definitions only superficially.

11.3 Conclusion

This thesis was comprised of one psychometric study and four experimental studies of current EI theory and the *awareness* and *flexibility* conjecture. In each of the studies hypotheses were generated based on either of these two theories although in most of the studies, these hypotheses were not mutually exclusive. The conjecture received very limited support, as did the conjecture. An attempt could be made to 'explain away' these findings, but these claims are ultimately unconvincing. It would seem that these preliminary results do not support the retention of the conjecture.

References

Austin, E.J. (2004). An investigation of the relationship between trait emotional intelligence and emotional task performance. *Personality and Individual Differences, 36,* 1855-1864

Austin, E.J. (2005). Emotional intelligence and emotional information processing. *Personality and Individual Differences, 39,* 403-414.

Austin, E.J., Saklofske, D.H., Huang, S.H.S., & McKenney, D. (2004). Measurement of trait emotional intelligence: testing and cross-validating a modified version of Schutte et al.'s (1998) measure. *Personality and Individual Differences, 36,* 555-562.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-function Autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5-17.

Bar-on, R. (1997b). Bar-On emotional quotient inventory: User's manual. Toronto: Multihealth Systems.

Bar-On, R., Brown, J.M., Kirkcaldy, B.D., & Thome, E.P. (2000). Emotional expression and implications for occupational stress; an application of the Emotional Quotient Inventory (EQ-i). *Personality and Individual Differences, 28,* 1107-1118

Bastian, V.A., Burns, N.R., & Nettlebeck, T. (2005). Emotional intelligence predicts life skills, but not as well as personality and cognitive abilities. *Personality and Individual Differences,39,*1134-1145.

Boucsein, W. (1992). *Electrodermal Activity.*Pennsylvania: Plenum Press.

Brackett, M. A., & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin, 29,* 1147-1158.

Bradley, M. M. (2000). Emotion and motivation. In J.T. Cacioppo, L. G. Tassinary, and G. Berntson (Eds.) *Handbook of Psychophysiology* (pp. 602-642). New York: Cambridge University Press.

Butler, E.A., Egloff, B., Wilhelm, F.H., Smith, N.C., Erickson, E.A., & Gross, J.J. (2003). The social consequences of expressive suppression. *Emotion, 3,* 48-67.

Cattell, R. B., Baggaley, A.R. (1960). The salient similarity index for factor matching. *British Journal of Statistical Psychology,* 13, 33-46.

Ciarrochi, J.V., Chan, A.Y.C., & Caputi, P. (2000). A critical evaluation of the emotional intelligence construct. *Personality and Individual Differences, 28,* 539-561

Davies, M., Stankov, L., & Roberts, R.D., (1998). Emotional intelligence: In search of an elusive construct. *Journal of Personality and Social Psychology, 75,* 989-1015

Dawda, D., & Hart, S.D. (2000). Assessing emotional intelligence: reliability and validity of the Bar-On Emotional Quotient Inventory (EQ-i) in university students. *Personality and Individual Differences, 28,* 797-812

Dawson, M.E., Schell, A.M., & Filion D.L. (2000). The electrodermal system. In J.T. Cacioppo, G. Tassinary, and G. Berntson (Eds.) *Handbook of Psychophysiology* (pp. 602-642). New York: Cambridge University Press.

Day, A.L., & Carrol, S.A. (2004). Using an ability-based measure of emotional intelligence to predict individual performance, group performance, and group citizenship behaviours. *Personality and Individual Differences, 36,* 1443–1458

Deary, I. J. (2000) *Looking Down on Human Intelligence.* New York: Oxford University Press.

Deary, I., Egan, V., Gibson, G. J., Austin, E. J., Brand, C. R., & Kellaghan, T. (1996). Intelligence and the differentiation hypothesis. *Intelligence, 23*, 105-132.
Deiner, E., Larsen, R J., Levin, S., & Emmons, R.A. (1985). Intensity and frequency: dimensions underlying positive and negative affect.
De Raad, B., & Perugini, M. (Eds). *Big Five Assessment,* Seattle: Hogrefe & Huber.
Edelberg, R. (1993). Electrodermal mechanisms: a critique of the two-effector hypothesis and a proposed replacement. In J-C Roy, ., W. Boucsein, D.C. Fowles, & J.H. Gruzelier (Eds), *Progress in Electrodermal Research* (pp. 7-31). London: Plenum Press.
Elklit, A. (1996). Coping styles questionnaire: a contribution to the validation of a scale for measuring coping strategies. *Personality and Individual Differences, 21,* 809-812.
Ekman, P., & Friesen, W.V. (1996) A new pan-cultural facial expression of emotion. *Motivation and Emotion, 10,* 159-168
Engelberg, E., & Sjoberg, L. (2004). Emotional intelligence, affect intensity, and social adjustment. *Personality and Individual Differences, 37*, 533-542.
Farrelly, D., & Austin, E. J. (Submitted) Ability EI as an intelligence? A preliminary study of associations of the MSCEIT with performance on emotion-processing and social tasks and with cognitive ability. *Cognition and Emotion.*
Furnham, A., & Buchanan, T. (2005). Personality, gender and self-perceived intelligence *Personality and Individual Differences, 39,* 543-555.
Forgas, J.P. (1998). On being happy and mistaken: mood effects on the fundamental attribution error. *Journal of Personality and Social Psychology, 75,* 318-331
Forgas, J.P., & Cromer, M. (2004). On being sad and evasive: Affective influences on
verbal communication strategies in conflict situations *Journal of Experimental Social Psychology, 40,* 511-518
Fridlund, A. (1991) Evolution and facial action in reflex, social motive, and paralanguage. *Biological Psychology, 32,* 3-100
Gleitman, H., Fridlund, A. J., & Reisberg, D. (2004). *Psychology: Sixth Edition.* New York: W.W. Norton & Company.
Gardner, H. (1983) *Frames of Mind.* New York; Basic Books.
Gannon, N., & Ranzijn, R. (2005). Does emotional intelligence predict unique variance in life satisfaction beyond IQ and personality? *Personality and Individual Differences, 38,* 1354-1363.
Gignac, G.E., Palmer, B.R., Manocha, R., & Stough, C. (2005). An examination of the factor structure of the Schutte self-report emotional intelligence (SSREI) scale via confirmatory factor analysis. *Personality and Individual Differences, 39,* 1029-1042.
Gohm, C.L., Corser, G.C., Dalsky, D.J. (2005). Emotional intelligence under stress: Useful, unnecessary, or irrelevant? *Cognition and Emotion,*
Gould, S.J. (1981). *The Mismeasure of Man.* Harmondworth, UK: Pelican Books Ltd.
Gross, J. J. (1999). Emotion regulation: Past, present, future. *Cognition and Emotion, 13*, 551-573.
Gross, J.J., & John, O.P. (2000). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85,*348-362.

Hartmann, P., Teasdale, T., W. (2004). A test of Spearman's "Law of Diminishing Returns" in a large sample of Danish military draftees. *Intelligence, 32,* 499-508.
Hartmann, P., Teasdale, T., W. (2005). Spearman's "Law of Diminishing Returns" and the role of test reliability investigated in a large sample of Danish military draftees. *Personality and Individual Differences, 39,* 1193-1203
Jausovec, N., & Jausovec, K. (2005). Sex differences in brain activity related to general and emotional intelligence. *Brain and Cognition, 59,* 277-286.
Jensen, A. R. (2003). Regularities in Spearman's law of diminishing returns. *Intelligence, 31,* 95-105.
Jensen, A.R. (1981). *Straight Talk about Mental Tests.* London: Methuen & Co.
Jensen, A.R. (1980). *Bias in Mental Testing.* New York: Free Press.
Kafetsios, K. (2004). Attachment and emotional intelligence abilities across the life course. *Personality And Individual Differences, 37*(1), 129-145.
Kline, P. (1991). *Intelligence: The Psychometric View.* London: Routledge.
Lang, P.J., Bradley, M.M., & Cuthbert, B.N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual. Technical Report A-6.* University of Florida, Gainesville, FL.
Leible, T. L., & Snell, W. E. (2004). Borderline personality disorder and multiple aspects of emotional intelligence. *Personality and Individual Differences, 37*(2), 393-404.
Locke, E.A. (2005). Why emotional intelligence is an invalid concept. *Journal of Organizational Behavior,29,* 425-431
MacCann, C., Roberts, R.D., Matthews, G., & Zeidner, M. (2004). Consensus scoring and empirical option weighting of performance-based Emotional Intelligence (EI) tests. *Personality and Individual Differences, 36,* 645-662.
Martin, M. (1990). On the Induction of Mood. *Clinical Psychology Review, 10*(6), 669-697.
Matthews, G., Zeidner, M., & Roberts, R.D. (2002). *Emotional Intelligence: Science and Myth.* Cambridge, MA; MIT Press.
Mayer, J., Caruso, D., & Salovey, P. (1999). Emotional intelligence meets traditional for an intelligence. *Intelligence, 27,* 267-298
Mayer, J. D., Salovey, P., & Caruso, D. R. (2002). *Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT) user's manual.* Toronto, Ontario, Canada: MHS Publishers.
Mayer, J. D., Salovey, P., Caruso, D., & Sitarenios, G. (2003). Measuring Emotional Intelligence With the MSCEIT V2.0. *Emotion, 3*(1), 97-105.
Parker, J.D.A., Saklofske, D.H., Shaughnessy, P. A., Huang, S.H.S., Wood, L.M., & Eastabrook, J.M. (2005). Generalizability of the emotional intelligence construct: A cross-cultural study of North American aboriginal youth. *Personality and Individual Differences, 39,* 215–227.
O'Connor, R. M., & Little, I. S. (2003). Revisiting the predictive validity of emotional intelligence: self-report versus ability-based measures. *Personality and Individual Differences, 35,* 1893-1902.
Palmer, B.R., Gignac, G., Manocha, R, & Stough, C. (2005). A psychometric evaluation of the Mayer–Salovey–Caruso Emotional Intelligence Test Version 2.0. *Intelligence, 33,* 285-305.

Parker, J. D. A., Taylor, G. J., & Bagby, R. M. (2001). The relationship between emotional intelligence and alexithymia. *Personality and Individual Differences, 30*(1), 107-115.

Petrides, K. V. & Furnham, A. (in press). The role of trait emotional intelligence in a gender-specific model of organizational variables. *Journal of Applied Social Psychology*

Petrides, K.V., & Furnham, A. (2000). On the dimensional structure of emotional intelligence. *Personality and Individual Differences, 29,* 313-320.

Petrides, K.V., & Furnham, A. (2003). Trait emotional intelligence: behavioural validation in two studies of emotion recognition and reactivity to mood induction. *European Journal of Personality, 17,* 39-57.

Power, M., & Dalgleish, T. (1997). *Cognition and Emotion: From Order to Disorder.* Hove: Psychology Press.

Roger, D., Jarvis, G., & Najarian, B. (1993). Detachment and coping: the construction and validation of a new scale for measuring coping strategies. *Personality and Individual Difference,* 15, 619-626.

Rosenthal, R., & Rosnow, R.L.(1991). *Essentials of Behavioral Research: Methods and Data Analysis.* London: McGraw-Hill

Saklofske, D.H., Austin, E.J., & Minski, P.S. (2003). Factor structure and validity of a trait emotional intelligence measure. *Personality and Individual Differences, 34,* 707-721.

Saklofske, D.H., Austin, E.J., Galloway, J., & Davidson, K. (2007). Individual difference correlates of health-related behaviours: preliminary evidence for links between emotional intelligence and coping. *Personality and Individual Differences,* 42, 491-502.

Salovey, P., & Mayer, J. (1989/90). Emotional intelligence. *Imagination, Cognition, and Personality, 9,* 185-21

Schutte, N.S., Malouff, J.M., Hall, L.E., Haggerty, D.J., Cooper, J.T., Golden, C.J., & Dornheim, L. (1998). Development and validation of a measure of emotional intelligence. *Personality and Individual Differences, 25,* 167-177.

Scargle, J. D. (2000). Publication bias: the "file-drawer" problem in scientific inference. *Journal of Scientific Exploration, 14,* 91-106

Schachter, S., & Singer, J.E. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69,* 379-399.

Schutte, N.S., Malouff, J.M., Bobik, C., Coston, T.D., Greeson, C., Jedlicka, C., Rhodes, E., & Wendorf, G. (2001). Emotional intelligence and personal relations. *Journal of Social Psychology, 141,* 523-536

Schulte, M.J, Ree, M.J., & Carretta, T.R. (2004). Emotional intelligence: Not much more than g and personality. *Personality and Individual Differences, 39,* 1059-1068

Smith, E., Nolen-Hoeksema, S., Fredrickson, B., Loftus, G.R., Bem, D., & Maren, M. (2003). *Atkinson & Hilgard's Introduction to Psychology.* Belmont, CA: Thomson Wadsworth.

Snyder, M (1974). Self-monitoring of expressive behaviour. *Journal of Personality and Social Psychology, 30,* 526-537

Spearman, C. (1927). *The Abilities of Man.* London: Macmillan.

Strongman, K.T. (2003). *The Psychology of Emotion.* Chichester: W.J. Wiley & Sons.

Van Der Zee, K. & Wabeke, R. (2004). Is trait-emotional intelligence

simply or more than just a trait? *European Journal of Personality, 18,* 243–263
Van Rooy, D. L., & Viswesvaran, C. (2004). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal Of Vocational Behavior, 65*(1), 71-95.
Van Rooy, D.L., Alonso, A., & Viswesvaran, C. (2005). Group differences in emotional intelligence score: theoretical and practical implications. *Personality and Individual Differences, 38,* 689-700.
Warwick, J., & Nettelbeck, T. (2004). Emotional intelligence is.? *Personality and Individual Differences, 37*(5), 1091-1100.
Westermann, R., Spies, K., Stahl, G., & Hesse, F.W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology, 26,* 557-580.
Wong, C., & Law, K.S. (2002). The effects of leader and follower emotional intelligence on performance and attitude: An exploratory study. *The Leadership Quarterly, 13,* 243-274.
Zack, M., Toneatto, T., & Streiner, D. L. (1998). The SCL-90 factor structure in comorbid substance abusers. *Journal of Substance Abuse,* 10, 85-101.